

# Interacção 2010

Actas da 4ª Conferência Nacional em Interacção Humano-Computador

13, 14 e 15 de Outubro, 2010  
Aveiro, Portugal

## Organização e Apoios



## Patrocínios



Actas da 4ª Conferência Nacional em Interacção Humano-Computador  
Aveiro, 13, 14 e 15 de Outubro de 2010

Óscar Mealha  
Joaquim Madeira  
Daniel Tércio  
Beatriz Sousa Santos  
(Editores)

Capa e Logótipo: Pedro Amado

Edição  
Grupo Português de Computação Gráfica  
(Eurographics Portuguese Chapter)  
Edifício INESC Porto  
Campus da FEUP  
Rua Dr. Roberto Frias, 378  
4200-456 Porto, Portugal

ISBN: 978-972-98464-4-1

Depósito Legal: 316940/10

Tiragem: 100 exemplares

Impressão e Acabamentos: Designeed, Lda.

Outubro, 2010



## Prefácio

Encontramo-nos perante mais uma edição do congresso nacional sobre Interacção, o “Interacção 2010”, um evento científico cuja regularidade tem permitido consolidar a sua natureza profundamente transdisciplinar. A referência do congresso continua a ser a área científica internacional de “*Human-Computer Interaction*”, uma área que tem ganhado relevância e valor no progresso sócio-económico desta última década, onde o ser humano e a coisa social assumem a prioridade que lhes é devida.

O “Interacção” tem assumido sempre a condição da debater um vasto leque de temas, mas sempre em contextos de mediação tecnológica e onde o objectivo final é servir a condição humana. Esta edição continua a confirmar esta atitude que se reflecte nas grandes áreas temáticas que anunciou para submissão de trabalhos: Concepção, Desenvolvimento e Estudo da Interacção, Tecnologias e Aplicações de Suporte à Interacção, Aspectos Humanos, Sociais e Organizacionais, e Contextos de Interacção diversos.

A resposta da comunidade científica nacional ao apelo de participação foi de extrema eficiência e riqueza. Foram submetidos um total de 64 artigos longos, artigos curtos e posters e/ou demos. O resultado da avaliação (“*double-blind*”) pelos revisores, e as conversões de formato sugeridas a alguns autores, conduziram a que o leitor venha a encontrar nestas actas 21 artigos longos, 20 artigos curtos e 15 posters. Por insuficiência de propostas de demos esta possibilidade de exposição não foi considerada.

Esta quantidade crescente de trabalho em relação a edições anteriores do “Interacção”, e os três oradores convidados (2 internacionais e 1 nacional), originaram um programa de trabalho que incontornavelmente se estendeu a três dias de trabalho contrariando os tradicionais dois dias de congresso.

O programa decorreu de 13 a 15 de Outubro de 2010 na Universidade de Aveiro e as 8 sessões que o caracterizaram foram organizadas respeitando os temas: i) Interfaces Multi-toque; ii) Interacção com Dispositivos Móveis; iii) Concepção e Desenvolvimento da Interacção; iv) Realidade Virtual e Aumentada; v) Visualização de Informação; vi) Videojogos e Interacção; vii) Acessibilidade e viii) Interacção em Contextos Diversos.

Por fim, e porque um evento desta natureza não se consegue organizar sem a mobilização, o trabalho de grande qualidade e a dedicação de várias pessoas, importa agradecer a todos os elementos da Comissão Organizadora, e aos estudantes voluntários que ajudaram em toda a logística, bem como a todos os elementos da Comissão de Programa e aos revisores externos que também aceitaram analisar trabalhos submetidos.

Aos nossos convidados, Francisco Providência, Jennifer Preece e Shahram Izadi, pela disponibilidade de estarem e partilharem connosco o seu conhecimento e experiência na área, o nosso sincero obrigado.

Pelo papel fundamental de um secretariado profissional e eficiente fica ainda o nosso agradecimento à Anabela Viegas.

Terminamos com o especial agradecimento a todos os autores e conferencistas que acreditaram na competência e rigor de toda esta equipa para acolher o seu trabalho e expô-lo publicamente.

Despedimo-nos acreditando que o “Interacção” continuará por terras lusas, quiçá numa toada cada vez mais internacionalizada, e deixando votos para que a edição de 2012 seja pautada por igual sucesso e entusiasmo; a área da Interacção assim o exige e merece.

Bem hajam.

Aveiro, Outubro de 2010

Óscar Mealha  
Daniel Tércio  
Beatriz Sousa Santos  
Joaquim Madeira

Aveiro, 13, 14 e 15 de Outubro de 2010

## Comissão Organizadora

Beatriz Sousa Santos – IEETA/DETI-UA; Joaquim Madeira – IEETA/DETI-UA (Co-Presidentes)

Ana Veloso – CETAC.media/DeCA-UA  
 Leonor Teixeira – GOVCOPP/DEGEI-UA  
 Paulo Dias – IEETA/DETI-UA

Anabela Viegas – IEETA (Secretariado)

Pedro Amado - DeCA-UA  
 Ravi Passos - DeCA-UA  
 Samuel Almeida - DeCA-UA

## Comissão Científica

### Co-Presidentes

Óscar Mealha – CETAC.media/DeCA-UA; Daniel Tércio – FMH-UTL

### Comité de Programa

Adérito Marcos - U. Aberta  
 Alfredo Ferreira - INESC-ID  
 Ana Amélia Carvalho - DCTE/UM  
 Ana Cristina Ramada Paiva - FEUP  
 Ana Isabel Veloso - CETAC.media/DeCA-UA  
 Ana Paula Afonso - FCUL  
 Ana Paula Cláudio - FCUL  
 António Augusto de Sousa - FEUP  
 António Branco - FCUL  
 António Nestor Ribeiro - UM  
 António Ramires Fernandes - UM  
 Beatriz Sousa Santos – IEETA/DETI-UA  
 Carlos Duarte - FCUL  
 Carlos J. Costa - ADETTI/ISCTE  
 Daniel Gonçalves - IST/UNL  
 Eduardo Dias - UE/Ydreams  
 Francisco Rebelo - FMH  
 Gonçalo Furtado - FAUP  
 Isabel Trancoso – IST-UTL  
 Javier Rodeiro Iglesias – Univ de Vigo  
 Joaquim Madeira – IEETA/DETI-UA  
 João Brisson Lopes – IST-UTL  
 João Duarte Cunha - LNEC  
 João R. Caldas Pinto – IST-UTL  
 Joaquim Jorge – INESC/IST-UTL  
 Jorge A. Santos - UM  
 Jorge Ferraz de Abreu - CETAC.media/DeCA-UA  
 José Campos - IPV  
 José Creissac Campos - DI/UM  
 Leonor Teixeira – GOVCOPP/DEGEI-UA

Luís Borges Gouveia - UFP  
 Licínio Roque - UC  
 Lúcia Oliveira - CETAC.media/DeCA-UA  
 Luís Carrigo - FCUL  
 Manuel João Fonseca - INESC-ID  
 Manuel Próspero dos Santos - FCT/UNL  
 Margarida Almeida - CETAC.media/DeCA-UA  
 Maria Beatriz Carmo – FCUL  
 Mário Rui Gomes – IST-UTL  
 Miguel Sales Dias - ISCTE/Microsoft  
 Nathalie Aquino - CIMPS / Un. Politécnica de Valencia  
 Nuno Correia - FCT/UNL  
 Nuno Guimarães- FCUL  
 Nuno Jardim Nunes - UMa  
 Nuno Otero - UM  
 Paula Alexandra Silva - Fraunhofer Portugal/FEUP  
 Paulo Dias - IEETA/DETI-UA  
 Pedro Almeida - CETAC.media/DeCA-UA  
 Pedro Arezes - UM  
 Pedro Faria Lopes - ISCTE  
 Rafael Bidarra - Delft Univ. of Technology - Holanda  
 Raquel Santos - FMH  
 Rui José - UM  
 Rui Raposo - CETAC.media/DeCA-UA  
 Sílvia Abrahão - Univ. Politécnica de Valência - Espanha  
 Teresa Chambel - FCUL  
 Teresa Romão - FCT/UNL

### Revisores adicionais

Luís Duarte, DI/ FCUL  
 José Moreira, IEETA/ UA

Paulo Pombinho, LASIGE/ FCUL  
 Rogério Bandeira, LASIGE/ FCUL

Aveiro, 13, 14 e 15 de Outubro de 2010

# Conteúdos

## ORADORES CONVIDADOS

<b>Making Computers More Natural to Use</b> <i>Shahram Izadi</i> .....	3
<b>Opportunities and Challenges for Technology Mediated Social Participation</b> <i>Jennifer Preece</i> .....	5
<b>Design de dispositivos de comunicação e interacção na museografia contemporânea</b> <i>Francisco Providência</i> .....	7

## SESSÕES TÉCNICAS

### Interfaces Multi-toque

<b>Conjuntos de gestos de comando para ferramentas de desenho em dispositivos sem teclado</b> <i>Tiago Gomes, Carlos Duarte, Luís Carriço, Joana Neca e Tiago Reis</i> .....	13
<b>Técnicas de interacção para revisão de cenários 3D</b> <i>Bruno de Araújo, Diogo Mariano, Ricardo Jota, Alfredo Ferreira e Joaquim Jorge</i> .....	21
<b>Interacção multi-toque no contexto do DJing</b> <i>Pedro A. Lopes, João A. Madeiras Pereira e Alfredo Ferreira</i> .....	29
<b>Combining different types of interaction on multi-touch surfaces</b> <i>Tarquínio Mota</i> .....	33

### Interacção com Dispositivos Móveis

<b>Designing interaction for outdoor sports performance analysis</b> <i>João M. Brisson Lopes</i> .....	39
<b>DETIGuide: Interagindo com um ecrã de grandes dimensões através de um dispositivo móvel</b> <i>Rui Palha, Paulo Dias, Rui A. Costa e Beatriz Sousa Santos</i> .....	47
<b>Utilização de dados fisiológicos na avaliação de aplicações móveis</b> <i>Luís Duarte, Marco de Sá e Luís Carriço</i> .....	55
<b>Automatic noise-based volume adjustments for mobile media players</b> <i>Tiago Reis, Luís Carriço e Carlos Duarte</i> .....	59

### Concepção e Desenvolvimento da Interacção

<b>The Genio centred on the human being — A methodology for fast development centred on the human being</b> <i>Duarte Teixeira e Mário Rui Gomes</i> .....	65
<b>Desenvolvimento centrado no utilizador em ambiente de <i>Participatory Design</i>: uma experiência na área da saúde</b> <i>Leonor Teixeira, Vasco Saavedra, Carlos Ferreira e Beatriz Sousa Santos</i> .....	73
<b>GUI Behavior from source code analysis</b> <i>João C. Silva, Carlos E. Silva, José C. Campos e João A. Saraiva</i> .....	81
<b>GUI for XML documents access using query-by-example paradigm</b> <i>Daniela Fonte, Daniela da Cruz, Pedro Rangel Henriques e Alda Lopes Gançarski</i> .....	89

## Realidade Virtual e Aumentada

<b>Crystal ball, virtual gallery creation system for immersive environments</b> <i>Rui Nuno de Almeida e Miguel Sales Dias</i> .....	97
<b>Aplicação de personagens virtuais como guias de visita</b> <i>Pedro Ribeiro, Tiago Silva e Rui José</i> .....	105
<b>Avaliação da interacção com envoltórios virtuais através de matrizes de exploração do espaço</b> <i>Emília Duarte, Júlia Teles, Luís Teixeira e Francisco Rebelo</i> .....	113
<b>Prototipagem rápida de ambientes ubíquos</b> <i>José Luís Silva, Óscar R. Ribeiro, João M. Fernandes, José Creissac Campos e Michael D. Harrison</i> .....	121
<b>Locomoção virtual via movimento físico do utilizador: controlo de velocidade do avatar</b> <i>Luís Bruno, Bruno de Araújo, João Pereira e Joaquim Jorge</i> .....	129
<b>Rato versus joystick como dispositivo de entrada para navegação num ambiente virtual com um Head-Mounted Device</b> <i>Paulo Dias, Paulo Santos, Carlos Ferreira e Beatriz Sousa Santos</i> .....	137
<b>On the impact of training HRTF-based auralisation</b> <i>Catarina Mendonça, Jorge A. Santos, Guilherme Campos, Paulo Dias e João P. Ferreira</i> .....	141

## Visualização de Informação

<b>Exploração e visualização de informação pessoal</b> <i>Paulo Gomes, Sandra Gama e Daniel Gonçalves</i> .....	147
<b>Metabrain — Conhecimento na era do PetaByte</b> <i>João Teixeira, Gabriel Barata e Daniel Gonçalves</i> .....	155
<b>Pontos de interesse relevantes num mapa: à procura de boas práticas</b> <i>Filipe Gil, Ana Paula Cláudio e Maria Beatriz Carmo</i> .....	159
<b>Pesquisa interactiva de fotografias com base na regra dos terços</b> <i>Gabriel Barata, Ricardo Dias, Sandra Gama, Manuel J. Fonseca e Daniel Gonçalves</i> .....	163
<b>Exploração interactiva de colecções musicais usando treemaps ordenados semanticamente</b> <i>Ricardo Dias e Manuel J. Fonseca</i> .....	167

## Videojogos e Interação

<b>Measuring difficulty in platform videogames</b> <i>Fausto José Mourato e Manuel Próspero dos Santos</i> .....	173
<b>Wiplay: avaliação da interacção</b> <i>André F. S. Barbosa e Frutuoso G. M. Silva</i> .....	181
<b>Augmented Reality without markers</b> <i>Luís F. A. Silva e Frutuoso G. M. Silva</i> .....	189
<b>2D game editor on-line — A cloud computing perspective on game edition</b> <i>Bruno Miguel Cardoso e Teresa Romão</i> .....	193
<b>MMORPGs e culturas de convergência</b> <i>Rui Raposo e Nídia Salomé Morais</i> .....	197

## Acessibilidade

<b>Caderno escolar electrónico adaptativo — Resultados preliminares</b> <i>Luís Alexandre e Salvador Abreu</i> .....	205
<b>Promovendo a colaboração entre crianças do pré-escolar através da Realidade Aumentada</b> <i>Sofia Pessanha e Pedro Campos</i> .....	213
<b>Rumo a interfaces tácteis acessíveis</b> <i>Tiago Guerreiro, Hugo Nicolau, Joaquim Jorge e Daniel Gonçalves</i> .....	221
<b>Interacção do sénior com jardim de lar: desenvolvimento de um guião de entrevista</b> <i>Cláudia Mourato Nunes, Francisco dos Santos Rebelo e Fernando Moreira da Silva</i> .....	229
<b>Reabilitar o processo de reabilitação</b> <i>Teresa Gama, Tiago Guerreiro, Hugo Nicolau e Joaquim Jorge</i> .....	237
<b>Bases para reabilitação assistida</b> <i>Rita Pereira, Tiago Guerreiro, Hugo Nicolau, Daniel Gonçalves e Joaquim Jorge</i> .....	241
<b>Psicoterapia georreferenciada com dispositivos móveis</b> <i>Marco de Sá, Luís Carriço, João Faria, Isabel Sá, Gustavo Zurita, Nelson Baloian e Ricardo Pereira</i> .....	245
<b>Terapia do medo para crianças: desenho e prototipagem de uma ferramenta para dispositivos móveis</b> <i>Marco de Sá, Luís Carriço, Pedro Bernardo, João Faria e Isabel Sá</i> .....	249

## Interacção em Contextos Diversos

<b>Estudo comparativo de aplicações para a construção de modelos LEGO</b> <i>Daniel Mendes e Alfredo Ferreira</i> .....	255
<b>Empirical study of user behavior on home page layouts: SAPO portal case study</b> <i>Samuel Almeida, Óscar Mealha, Ana Isabel Veloso e André Luís</i> .....	263
<b>Interfaces para Hipervídeo em 360°</b> <i>Luís A. R. Neng e Teresa Chambel</i> .....	271
<b>The last words of Domenico: um instrumento de performance para a reconstrução sonora de discursos</b> <i>Vitor Lago Silva e Pedro Branco</i> .....	275

## Posters

<b>Metodologias de análise visual baseadas na tecnologia de Realidade Virtual: Jardim Botânico do Rio de Janeiro, Brasil e Jardim do Cerco, Portugal</b> <i>Aurélio Nogueira, Teresa Heitor, Maria Carreira, Gelly Rodrigues e Ana Pedroso</i> .....	281
<b>Áudio Descrição na Televisão Digital Terrestre (TDT)</b> <i>Rita Oliveira, Jorge Ferraz de Abreu e Ana Margarida Almeida</i> .....	283
<b>Tree on Mars, an Immersive Virtual Reality Experience</b> <i>Paulo Ricardo Duarte e Miguel Sales Dias</i> .....	285
<b>Elaboração de Conteúdo 3D para utilização na CaveHollowspace do Lousal</b> <i>Luís Miguel S Ponciano e Miguel Sales Dias</i> .....	287
<b>O Tacto e a Interação Móvel</b> <i>João Benedito, Tiago Guerreiro, Hugo Nicolau Joaquim Jorge e Daniel Gonçalves</i> .....	289

<b>Cognimobile: Diferenças Cognitivas e os Dispositivos Móveis</b> <i>João Oliveira, Tiago Guerreiro, Hugo Nicolau e Daniel Gonçalves</i> .....	291
<b>ClimbOn – An infinite game with difficulty adjustment</b> <i>Ricardo Ameixa, Fausto Mourato e João Moraes</i> .....	293
<b>Sistema multi-modal de identificação de utilizadores IPTV- um processo de investigação</b> <i>Telmo Silva, Jorge Trinidad Ferraz de Abreu e Osvaldo Rocha Pacheco</i> .....	295
<b>Support for inferring user abilities for multimodal applications</b> <i>Carlos Duarte, Pedro Feiteira, Daniel Costa e David Costa</i> .....	297
<b>Avaliação pericial de barreiras ao acesso sobre sítios web de entidades públicas</b> <i>Nádia Fernandes, Rui Lopes e Luís Carriço</i> .....	299
<b>NavTilt: Interface Gestual para Cegos</b> <i>David Lucas, Hugo Nicolau, Tiago Guerreiro e Joaquim Jorge</i> .....	301
<b>Email Visualization: A Context-Based Approach</b> <i>Celso Cardoso e Paula Alexandra Silva</i> .....	303
<b>Time Automaton - A visual mechanism for temporal querying</b> <i>Luís Certo, Teresa G. Dias e José Borges</i> .....	305
<b>Um protótipo de um sistema de apoio à decisão em actividades logísticas de localização e distribuição</b> <i>Rui Borges Lopes, Carlos Ferreira e Beatriz Sousa Santos</i> .....	307
<b>m-Tourism 2.0 os sistemas integrados de multiagentes para apoio à informação e comunicação no turismo</b> <i>Pedro Beça e Rui Raposo</i> .....	309

## Índice de Autores

<b>Autores</b> .....	313
----------------------	-----



## **Oradores Convidados**

Aveiro, 13, 14 e 15 de Outubro de 2010

# Making Computers More Natural to Use

Shahram Izadi

Microsoft Research Cambridge  
7 JJ Thomson Avenue  
Cambridge, CB3 0FB, UK  
shahrami@microsoft.com

---

## Abstract

*The computing mouse has achieved a great deal for a device that's conceptually over four decades old. Indeed it is mainly because of this device that we have the Graphical User Interface (GUI), comprising windows, icons, menus, and pointer (WIMP). The mouse and WIMP are an integral part of our daily interactions with computers, but what's next? In this talk, I will give examples of novel computing devices being built at Microsoft Research, which allow users to shift away from the traditional mouse and WIMP-based interactions. Like other researchers, we are interested in enabling more natural interactions with computers, replacing the mouse with our hands, and making the user interface "come to life" in more tangible ways. Future interactions with computers are likely to become more hands-on, more playful, and more aligned with our real-world interactions.*

## Short Bio

SHAHRAM IZADI works at Microsoft Research UK, joining the company over 5 years ago. His research focuses on creating novel technologies that push the boundaries of how people interact with computers. In recent years, his research has focused on developing novel touch and gesture-based devices and interactive software. Thanks to his diverse computer science skills, he enjoys engaging in all areas of research, from the development of new hardware platforms right through to the design, implementation, and evaluation of the user interaction techniques. He has worked as a researcher at PARC, and was a research fellow on the Equator UK project, working at Nottingham University, University College London (UCL), and the Royal College of Art. He holds a PhD, masters and a degree (all with Honours) in Computer Science. His research has led to more than 50 patents and over 50 academic papers, including papers at UIST, CHI, and Ubicomp. He was awarded a best paper award for three consecutive years at UIST, and also holds paper awards at CHI, CSCW and Ubicomp. He has organized two international conferences and numerous workshops, and served as technical program chair and on various conference committees. He is currently engaged with various NUI-based tech-transfers within Microsoft. He was recently listed on the TR35, a list of 35 innovators, under the age of 35 who have made significant contribution to technology and the sciences.

Aveiro, 13, 14 e 15 de Outubro de 2010

# Opportunities and Challenges for Technology Mediated Social Participation

Jennifer J. Preece

College of Information Studies  
University of Maryland  
College Park, MD  
preece@umd.edu

---

## Abstract

*Technology-mediated social participation is generated when social networking tools (such as Facebook), blogs and microblogs (Twitter), user-generated content sites (YouTube), discussion groups, problem reporting, recommendation systems, and other social media are applied socially and for supporting national and international priorities such as health, energy, education, disaster response, environmental protection, conservation or community safety. There are opportunities for more strongly impacting national and international priorities by better understanding what motivates people's participation. At the same time we need to be aware of the dangers, including use by criminals and the need to protect personal privacy.*

## Short Bio

JENNIFER PREECE (<http://ischool.umd.edu/people/preece/>) is Professor and Dean in the College of Information Studies – Maryland's iSchool - at the University of Maryland. Jennifer's research focuses at the intersection of information, community and technology. She is particularly interested in community participation on- and off-line. She has researched ways to support empathy and social support online, patterns of online participation, reasons for not participating (i.e., lurking), strategies for supporting online communication, development of norms, and the attributes of successful technology-supported communities. Jennifer is author of over two hundred articles and eight books. Two recent books are: "Online Communities: Designing Usability, Supporting Sociability" (2000) and a co-authored best selling text entitled "Interaction Design: Beyond Human-Computer Interaction" (1<sup>st</sup> Ed. 2002; 2<sup>nd</sup> Ed. 2007; 3<sup>rd</sup> Ed. 2011) [www.id-book.com](http://www.id-book.com). Both books are published by John Wiley & Sons.

Aveiro, 13, 14 e 15 de Outubro de 2010

# Design de dispositivos de comunicação e interacção na museografia contemporânea

Francisco Providência

Departamento de Comunicação e Arte  
Universidade de Aveiro  
fprovidencia@ua.pt

---

## Abstract

*O design tem contribuído para a cultura material, através do desenho de artefactos, dispositivos e serviços de intermediação cultural. Se a escola de Ulm (Escola Superior da Forma de Ulm), constituiu a primeira experiência formativa do design apoiada na integração do conhecimento científico (T. Maldonado), por oposição e complementaridade com a sua antecessora Bauhaus, originalmente centrada no conhecimento artístico, os resultados da sua prática afirmaram-se universalmente por um design ergonomicamente destinado a funcionar, caracterizado por uma retórica estética da eficiência.*

*A forma das coisas, longe do tempo em que servia para encobrir a obsolescência comercial dos produtos, é o seu conteúdo de verdade (T. Adorno). A “forma” é o conteúdo, como o “meio” era já a mensagem (Mcluhan). Nesse sentido, se a tecnologia nunca foi neutra, então os seus objectivos deverão ser ponderados quanto à sua finalidade e interesse para o homem, já que, a sua adopção nem sempre resultou numa contribuição para o desenvolvimento humano, deixando antever a difícil compatibilização entre conforto e liberdade. Questionado o relacionamento do homem com a máquina feita à sua medida e imitação, observa-se que na essência da máquina encontramos o homem, mas na essência do homem não há nada de mecânico (M. Heidegger).*

*Ao projectar o artefacto, o designer questiona o mundo contemporâneo, produzindo a sua própria síntese, quer no desenho dos dispositivos, quer na assunção da sua responsabilidade ética, poeticamente veiculada. É na urgência de um “Design lacónico”, contribuinte para um mundo menos cínico, que o autor encontra, pela minimização da forma, as condições produtivas para que os artefactos não se substituam aos seus utilizadores. Objectos que parecem querer reencontrar um nível zero do entendimento, a tábua rasa da linguagem, o vazio de todo o sistema, a partir do qual se possa criar uma nova ecologia humana, mais verdadeira e vital.*

*É intenção desta comunicação mostrar exemplos da integração pelo Design, de tecnologias de comunicação e interacção, nos equipamentos museológicos de: Centro de Ciência Viva de Vila do Conde; Museu Municipal de Penafiel; Museu do Dinheiro / Banco de Portugal (projecto).*

## Short Bio

FRANCISCO PROVIDÊNCIA nasceu em Coimbra, em 1961, e formou-se em Design de Comunicação na Faculdade de Belas Artes da Universidade do Porto em 1985. Tem-se dedicado à actividade docente de fotografia, desenho na Universidade do Porto; dirige o Mestrado em Design na Universidade de Aveiro onde é Professor Associado Convidado. Com atelier próprio desde 1985, foi distinguido em 1999 com o Prémio Nacional de Design nas áreas da Comunicação, do Produto e do Ambiente, pelo Centro Português de Design, onde passou a colaborar como consultor. Foi premiado com a distinção Red dot em 2008.

Aveiro, 13, 14 e 15 de Outubro de 2010



## **Sessões Técnicas**

Aveiro, 13, 14 e 15 de Outubro de 2010

## **Sessão I**

### **Interfaces Multi-toque**

Aveiro, 13, 14 e 15 de Outubro de 2010

# Conjuntos de Gestos de Comando para Ferramentas de Desenho em Dispositivos sem Teclado

Tiago Gomes

Carlos Duarte

Luis Carriço

Joana Neca

Tiago Reis

LaSIGE, Faculdade de Ciências da Universidade de Lisboa

Campo Grande, Lisboa, Portugal

{tagomes,treis}@lasige.di.fc.ul.pt, {cad,lmc,jneca}@di.fc.ul.pt

## Sumário

*A interacção gestual, sem teclado, em aplicações de desenho, apresenta restrições mais rígidas do que em outro género de aplicações baseadas em gestos. Tal acontece, principalmente, pela necessidade de eliminar a ambiguidade existente quando um gesto representa um comando ou faz, simplesmente, parte do que está a ser desenhado. Neste artigo, exploramos conjuntos de gestos alternativos capazes de lidar com este problema. O contexto é fornecido através do ponto inicial do gesto ou, em alternativa, através de múltiplos toques no ecrã. Este artigo apresenta os resultados de uma experiência concebida para comparar dois conjuntos de gestos.*

## Palavras-chave

*Interacção gestual, Ferramentas de desenho, Comandos gestuais, Avaliação, Interacção sem teclado, Estudos de Utilizadores*

## 1. INTRODUÇÃO

Nas interfaces gestuais, os gestos são distinguidos, principalmente, pela forma que os descreve, pelo contexto final e pelo contexto inicial [Zelevnik06]. Nas ferramentas típicas de exploração/navegação, os gestos começam numa “página” e formam uma linha para a esquerda caso o desejado seja a “página seguinte”, ou para a direita caso seja a “página anterior”. Um gesto circular tipicamente significa a rotação da página. Se o gesto começar num objecto é expectável que os mesmos gestos movam o objecto. Contudo, se a aplicação é uma ferramenta de desenho, estes mesmos gestos são usados para desenhar linhas e círculos. O contexto inicial e final dificilmente resolve o problema, uma vez que alguém poderá querer desenhar em cima de desenhos já existentes.

Tipicamente, em ferramentas de desenho que requerem a utilização de rato existem duas soluções para este problema: 1) utilizar uma região do ecrã, por exemplo um menu ou paleta; 2) usar botões físicos auxiliares, tais como o botão direito do rato ou as teclas de *control* e *shift*. A primeira solução é fácil de lembrar, tem um nível de reconhecimento bastante preciso, mas é mais lenta [Balakrishnan04] [MacKenzie92]. Dependendo da alternativa, as opções clássicas baseadas em região requerem: a) um esforço extra para completar a tarefa e espaço extra no ecrã; b) um atalho complementar para *pop-ups* (e.g. botão direito do rato, manter uma tecla pressionada); c) em qualquer dos casos é necessário tempo extra para localizar a opção. Esta solução é frequentemente usada por novatos ou para comandos menos frequentes [Lane05].

Os botões físicos ou teclas são mais rápidos de utilizar. São usados como atalhos para executar comandos atómicos (e.g. *undo*, *redo*), comandos aplicados a argumentos previamente definidos (e.g. *delete*), ou para entrar em modo especial (e.g. agarrar para deslocar, mudar para ampliar). Para eliminar a ambiguidade entre desenho e comando, também se pode manter uma, ou mais, teclas pressionadas enquanto se executa o gesto.

Contudo, a utilização de dispositivos sem teclado ou com ecrãs de reduzidas dimensões, e nos quais o principal mecanismo de interacção é um dedo ou uma caneta/estilete, torna as coisas mais complicadas. Mesmo com o advento dos ecrãs multi-toque [Do-Lenh09], o desenho pode ser executado com mais que um dedo ao mesmo tempo, impossibilitando assim a utilização de gestos multi-toque como forma de eliminar a ambiguidade neste género de aplicações. Para além disto, desenhos mais precisos provavelmente serão executados com caneta/estilete, o que requer a utilização de uma mão para manusear esse periférico, enquanto a outra mão segura no dispositivo. Em dispositivos maiores, como os *tablets*, normalmente existe um botão extra na caneta, no entanto esse botão é culturalmente associado a um menu *pop-up*. Além disso, os dispositivos pequenos vêm equipados com estiletes sem botões extras. Por todas estas razões é necessário encontrar soluções para combater a ambiguidade.

Neste trabalho são propostas duas dimensões para distinguir gestos de comando de gestos de desenho em ferramentas de desenho para dispositivos sem teclado, baseados em toque. As dimensões surgiram da observação das actuais ferramentas de desenho, da literatura [Appert09]

[Bragdon09] [Zelevnik06] e de uma classificação sistemática de gestos sobre a possibilidade de discriminar características dos traços de desenho.

Gestos para alguns dos comandos mais comuns foram desenvolvidos e adaptados para o quadro conceptual de cada dimensão, tendo sido realizada uma experiência para comparar estes dois conjuntos. Um terceiro conjunto de controlo foi usado para avaliar a dificuldade adicional introduzida por estes dois conjuntos em comparação com uma situação ideal. No conjunto de controlo, a ferramenta foi capaz de adivinhar com 100% de precisão se o gesto era um comando, mesmo este sendo igual a um gesto de desenho (sem característica discriminatória).

As principais contribuições deste trabalho são: a) um par de dimensões sistemáticas que podem ser exploradas em aplicações com as características acima descritas; b) uma avaliação do preço que o utilizador está disposto a pagar para ter nestas ferramentas atalhos baseados em gestos.

A próxima secção apresenta uma revisão do trabalho relacionado. Seguidamente apresentamos a principal motivação para os dois conjuntos de gestos propostos. A secção seguinte detalha toda a experiência para avaliar a adequabilidade dos gestos propostos para aplicações com as características acima descritas. Posteriormente, é feita a análise dos dados recolhidos utilizando medidas qualitativas e quantitativas. Concluimos com um sumário das contribuições e uma breve apresentação do trabalho futuro a desenvolver.

## 2. TRABALHO RELACIONADO

A utilização de gestos através de rato, caneta ou estilete tem sido mapeada em diferentes tipos de comandos, quer seja para navegação em ‘web browsers’ [Moyle03], aplicações [Wu03] ou documentos [Smith04]. Contudo, para ferramentas de desenho, os gestos adoptados colidem com as normais actividades de desenho. Normalmente a solução adoptada é entrar em modo de comando, preferencialmente de uma forma temporária até que o comando seja realizado.

A utilização de gestos também é amplamente discutida para a selecção de itens em menus, exemplos desta discussão são encontrados em trabalhos como [Kurtenbach94] [Fitzmaurice03] [Guimbretière00] [Hinckley05] [Bau08]. Geralmente, estes sistemas proporcionam mecanismos para aceder a comandos através de menus radiais, organizados de maneira a otimizar a performance do utilizador. São particularmente interessantes para aceder facilmente a comandos quer numa fase de aprendizagem, onde o número de comandos é elevado, quer para comandos que são menos utilizados. Essencialmente estas aproximações diferem da nossa porque nós procuramos um subconjunto de comandos que sejam usados frequentemente e acessíveis através de teclas de atalho, constituindo assim um mecanismo complementar para aceder a um grande conjunto de comandos numa aplicação complexa.

Outra alternativa é a utilização de outras modalidades para complementar o gesto. Por exemplo o *QuickSet* [Cohen97] usa a fala para identificar um gesto de coman-

do, o que requer um esforço extra em reconhecimento de fala e dificilmente será funcional em cenários muito ruidosos. O *Silk* [Landay96] e o *Flatland* [Mynatt99] requerem que o utilizador pressione um botão para entrar em modo de comando. Parecendo uma boa e simples solução apresenta um problema crítico já referido anteriormente, não funciona para dispositivos sem botões físicos, a não ser que sejam usados botões virtuais causando assim um desperdício de espaço de ecrã.

Alguns trabalhos têm sido realizados para resolver as ambiguidades entre gestos de desenho livre e gestos de comando. Uma aproximação possível é a utilização de técnicas de mediação [Mankoff07]. Estas técnicas usam uma de duas estratégias: repetição ou escolha. Esta poderá ser uma boa solução mas o nosso objectivo é evitar quaisquer ambiguidades em vez de resolvê-las. Saund e Lank [Saund03] ofereceram uma solução para o problema do modo de comando baseada nas propriedades da trajectória da caneta e no contexto dessa mesma trajectória. Se esta diferenciação falhasse era utilizado um mediador de escolha sobre a forma de *pop-up*.

Outra forma possível de diferenciação é a utilização de uma marca posterior ao gesto, tendo sido já testada em alguns trabalhos. O *MathPad* [LaViola04] é um protótipo de uma aplicação matemática, a qual faz uso de um toque no fim de alguns gestos com o intuito de os distinguir. No *Tivoli* [Pedersen93] foi feita uma tentativa de usar um toque duplo como um indicador posterior, mas devido à dificuldade de o distinguir de um pequeno traço de desenho, esta abordagem foi abandonada e adoptaram um botão de caneta como indicador de gesto.

Um trabalho próximo do nosso é o *Fluid Inking* [Zelevnik06]. Zelevnik e Miller propuseram também uma aproximação onde a distinção entre gestos é definida, principalmente, no fim do gesto. Eles exploraram 3 formas diferentes de pontuação terminal: pausa, toque e utilização de um botão. A maior diferença entre os nossos trabalhos é que nós realizámos uma avaliação empírica comparativa e distinguimos os gestos somente pelo início dos mesmos. Uma análise comparativa de ambas as soluções poderá ser bastante interessante e alvo de trabalho futuro.

## 3. DESENVOLVIMENTO DOS CONJUNTOS DE GESTOS

Com o objectivo de identificar dimensões gestuais interessantes e relevantes, e que permitam eliminar a ambiguidade existente entre o gesto de desenho e o gesto de comando, analisámos aprofundadamente o trabalho já realizado [Appert09] [Bragdon09] [Duarte09] [Bau08] [Zelevnik06]. É de salientar que o foco não se encontra no gesto propriamente dito, mas sim na forma como os modos são identificados ou como é feita a troca dos mesmos. Além disso, assumimos a inexistência de teclas extra, quer no dispositivo em si (e.g. teclado) quer no periférico de interacção (e.g. dedo, caneta ou estilete).

Duas dimensões iniciais foram identificadas, abordando o início do gesto, nomeadamente: 1) a utilização de uma marca; 2) uma região específica do ecrã. No primeiro

caso a marca pode consistir num gesto muito pouco provável (e.g. um botão muito pequeno antes de iniciar o gesto) ou num toque duplo no ecrã antes de executar o gesto.

A utilização de uma zona dedicada do ecrã, segundo caso, poderá introduzir um problema de desperdício de espaço, assim como em ecrãs de maiores dimensões poderá levar a um esforço motor mais acentuado. O primeiro problema poderá ser minimizado através da utilização de uma pequena moldura invisível em redor do ecrã ou, caso seja necessário, só de um ou dois lados do ecrã. A segunda questão só se coloca para dispositivos de grandes dimensões, sendo o esforço motor desprezível para ecrãs de pequenas dimensões. Convém salientar que, em ambos os casos, os gestos de comando são tidos como atalhos que puderam ser complementados com a utilização de menus disponíveis através dos mecanismos clássicos.

A tabela 1 mostra os gestos definidos para alguns dos comandos mais utilizados em ferramentas de desenho: seleccionar, cortar e colar, apagar, *undo* e *redo*, e *scroll*. Na primeira coluna os gestos estão representados num contexto sem restrições, ou seja, assumindo que o sistema tem a capacidade de distinguir os gestos de comando dos gestos similares de desenho somente através da intenção do utilizador. Os gestos propriamente ditos são comuns a outras aplicações e foram considerados razoavelmente aceites em estudos relacionados [Neto09] [Bragdon09]. Os pequenos círculos representam objectos previamente desenhados, alguns desses objectos têm associada a simbologia de selecção, indicando assim que os comandos (apagar e cortar) são aplicados aos mesmos. Os comandos colar e *redo* utilizam gestos simétricos aos de copiar e *undo*. As setas existentes no desenho dos gestos indicam a direcção pelo qual o mesmo deve ser executado.

As outras duas colunas representam os mesmos gestos mas usando uma das dimensões acima discriminadas. Na coluna do meio, o círculo preenchido no início do gesto representa a marca que indica que o gesto será um comando. A escolha recaiu num toque duplo uma vez que este método de interacção é comum nas actuais interfaces de interacção com o utilizador. A última coluna representa a utilização da moldura invisível em redor do ecrã. Esta moldura está indicada por um rectângulo sombreado onde qualquer gesto de comando deve ser iniciado.

A utilização destes três conjuntos de gestos permite-nos compreender:

- o quão difíceis os utilizadores consideram os dois conjuntos pragmáticos em comparação com a situação ideal;
- As diferenças entre os dois conjuntos de gestos propostos;

#### 4. COMPARAÇÃO ENTRE OS CONJUNTOS DE GESTOS

Tendo como objectivo comparar os dois conjuntos de gestos anteriormente mencionados, preparámos uma

experiência para determinar qual dos conjuntos é mais apropriado para as condições descritas.

	Unrestricted	Depending on starting mark (e.g. d-click)	Depending on starting region (e.g. frame)
Scroll			
Select			
Copy/Paste			
Undo/Redo			
Delete			

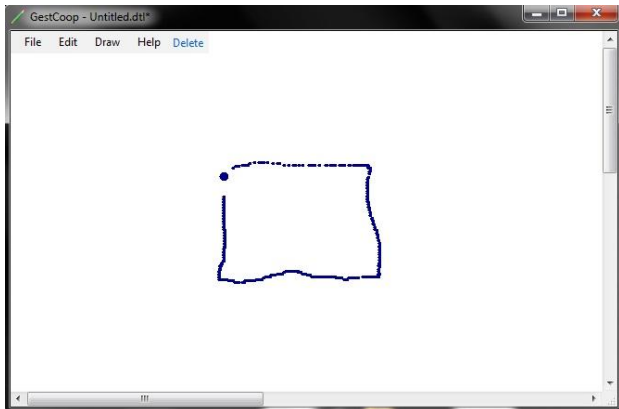
**Tabela 1: Gestos para os três conjuntos de gestos**

Nesta experiência os utilizadores tiveram de desempenhar tarefas numa aplicação de desenho simplificada, sendo que para o fazerem somente puderam recorrer a gestos. A aplicação suportava o desenho de duas formas geométricas: elipses e rectângulos (figuras 1 e 2). Além de desenhar as formas os utilizadores podiam arrastá-las pela área de trabalho e aplicar-lhes comandos como copiar, colar e apagar (figura 3). Também era possível fazer *undo* e *redo* às acções realizadas. A selecção de uma forma era realizada tocando sobre a mesma, e a selecção de múltiplas formas era possível através do gesto de selecção (figura 4). Sendo a área de trabalho maior que a janela de visualização, os utilizadores também podiam executar gestos com o objectivo de fazer *scroll* (as barras de scroll encontravam-se desactivadas, funcionando apenas como um mecanismo de feedback).

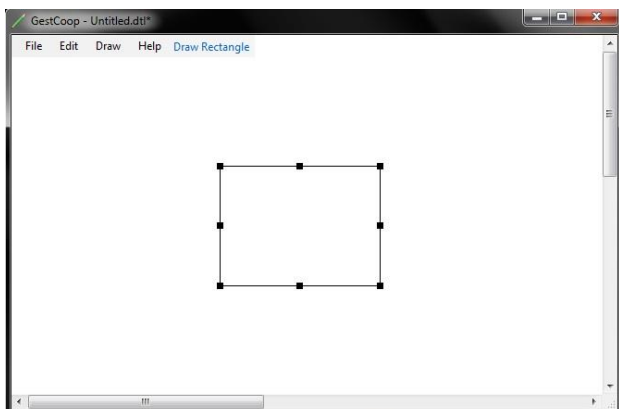
Querendo eliminar a influência de variáveis externas nesta avaliação, optámos por realizar a experiência como Feiticeiro de Oz. Assim, reduzimos o impacto que o reconhecimento dos gestos poderia ter nos resultados comparativos. Tomada esta opção, só os gestos de elipse, rectângulo e selecção eram reconhecidos automaticamente pela aplicação, sendo os restantes gestos reconhecidos por um dos avaliadores que monitorizavam a experiência.

Para avaliar a adequabilidade dos dois conjuntos de gestos, medimos ambos com métricas quantitativas e qualitativas. As métricas quantitativas incluem o tempo médio que os utilizadores demoraram a efectuar os diferentes gestos e uma métrica da qualidade do gesto. Esta métrica foi calculada por outra revisão de um avaliador, classificando todos os gestos realizados por todos os participantes nos testes. Todos os gestos foram capturados durante a experiência, e com base nesta captura, o avaliador clas-

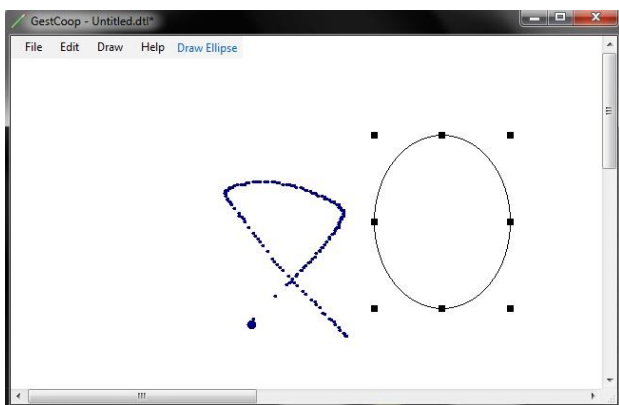
sificou os gestos dentro de uma das acções que eram possíveis reconhecer. Um gesto era considerado de qualidade se este reconhecimento coincidia com a intenção do utilizador ao realizar o gesto. As métricas qualitativas incluem uma avaliação subjectiva dos utilizadores aos três conjuntos de gestos, sendo que as características avaliadas foram a velocidade, o conforto, a precisão e a satisfação de utilização.



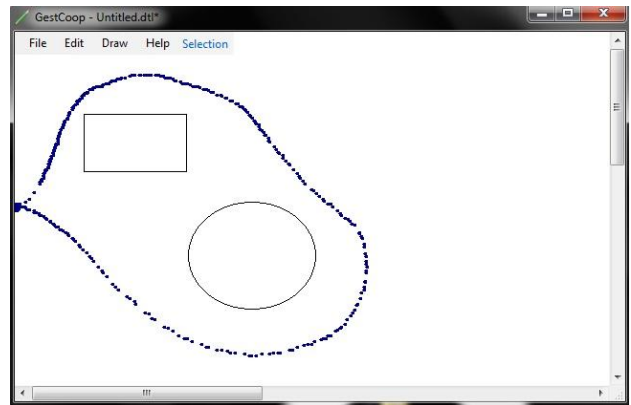
**Figura 1: Rectângulo a ser desenhado**



**Figura 2: Rectângulo depois de ser reconhecido**



**Figura 3: Apagando uma elipse em modo duplo toque**



**Figura 4: Selecção no modo moldura invisível**

#### 4.1 Procedimento

A experiência desenrolou-se em duas sessões com um dia de intervalo, tendo sempre os mesmos participantes.

A primeira sessão começou com uma explicação dos objectivos aos participantes, os quais seguidamente preencheram um questionário de perfil. Posteriormente, e para o primeiro conjunto de gestos, uma folha com ilustrações dos gestos foi fornecida aos participantes para que os mesmos se adaptassem à aplicação, tendo sempre o apoio de um dos avaliadores. Quando se sentiram preparados, a folha foi retirada e foi pedido aos participantes que desempenhassem um conjunto de tarefas pensadas para englobarem cada uma das acções pelo menos uma vez. Após todas as tarefas concluídas o segundo conjunto de gestos foi apresentado e todo o procedimento repetido, repetindo-se mais uma vez para o último conjunto de gestos. No final foi pedido aos participantes que preenchessem um questionário de satisfação. Todos os gestos realizados durante a experiência foram capturados através de *snapshots*.

Durante a segunda sessão nenhuma folha com ilustrações dos gestos esteve disponível. A sessão começou com uma breve descrição da aplicação, mas sem nunca mencionar os três conjuntos de gestos. Foi então pedido aos participantes que executassem um conjunto de tarefas, sem serem dadas quaisquer indicações sobre o conjunto de gestos a utilizar. Não existindo restrições conseguimos inferir qual o conjunto de gestos mais natural e intuitivo. Após completadas estas tarefas foi então pedido aos participantes que executassem um novo conjunto de tarefas mas desta vez repetindo-se para cada conjunto de gestos. Isto permitiu-nos detectar se os participantes tinham mais dificuldades em lembrarem os gestos de algum conjunto específico. Mais uma vez, e após terminarem todas as tarefas para todos os conjuntos de gestos, foi pedido aos participantes que preenchessem um novo questionário de avaliação. Também nesta segunda sessão todos os gestos foram capturados através de *snapshots*.

Sendo os gestos semelhantes, e não querendo que os resultados fossem afectados pelo efeito de treino, a ordem pelo qual os testes foram realizados por cada participante e para cada conjunto de gestos foi controlada em ambas as sessões.



## 4.2 Configuração da experiência

Para executar o procedimento descrito utilizámos um Tablet PC, nomeadamente um LG P100 com 10.6" de ecrã e com uma resolução de 1280 por 768 pixels. O computador encontrava-se configurado como um *tablet* com o ecrã orientado para o participante e o teclado coberto pelo mesmo. A aplicação de desenho encontrava-se maximizada ocupando todo o ecrã. Os participantes do teste encontravam-se sentados e a sua interacção com o tablet foi feita usando um estilete.

O avaliador que interpretava o papel de Feiticeiro de Oz encontrava-se perto do participante e com excelente visibilidade para o ecrã do Tablet PC. O reconhecimento dos gestos foi feito recorrendo a um teclado USB ligado ao Tablet PC. Este reconhecimento foi possível mapeando os gestos da aplicação para atalhos de teclado. Um outro avaliador estava presente na sala a realizar o papel de facilitador, cuja função foi descrever a experiência, fornecer ou retirar a folha com as ilustrações dos gestos e dar as instruções necessárias.

## 5. RESULTADOS

Um total de seis voluntários participou nas duas sessões experimentais. Com idades compreendidas entre os 23 e os 29 anos, a média de idades foi de 26 anos. Todos os participantes eram estudantes pós-graduados de Engenharia Informática e, segundo os mesmos, com mais de 10 anos de experiência de utilização de computadores. Cinco afirmaram conhecer pelo menos um dispositivo de interacção gestual, três dos quais com utilização regular. Esta utilização referia-se aos dispositivos Apple iPhone e iPod e Nintendo Wii.

Devido a um problema com a aplicação de *snapshots* nem todos os gestos da primeira sessão foram capturados, mesmo assim o total de gestos capturados durante as duas sessões ascendeu aos 986. A tabela 2 apresenta a distribuição dos gestos pelas acções efectuadas. Podemos observar que cerca de um terço dos gestos foram efectuados para comandos de *scroll*, e aproximadamente outro terço para comandos de apagar e de *undo*. O grande número de comandos de *scroll* deveu-se à natureza das tarefas a realizar, que implicavam muitas vezes a deslocação da janela de visualização ao longo da área de trabalho. O número de gestos de apagar e de *undo* ficou um pouco acima do esperado, principalmente devido a alguns erros de reconhecimento durante a acção de selecção. O problema subjacente a estes erros de reconhecimento será discutido mais à frente neste artigo.

Rectângulo	1.8%	Copiar	3.3%
Elipse	3.3%	Colar	3.3%
Arrastar	13.2%	<i>Undo</i>	11.7%
<i>Scroll</i>	34.9%	<i>Redo</i>	3.3%
Seleccção	12.5%	Apagar	12.6%

**Tabela 2: Distribuição dos comandos ao longo da experiência**

A tabela 3 apresenta a distribuição dos gestos pelos conjuntos de gestos. Quase metade dos gestos foram efectuados no conjunto sem restrições. Isto deveu-se ao grande número de tarefas efectuadas recorrendo a este conjunto durante a segunda sessão de testes, uma vez que numa parte da sessão não existia imposição quanto ao conjunto de gestos a utilizar e todos os participantes optaram pelo conjunto sem restrições.

Sem Restrições	Moldura	Duplo Toque
47.4%	29.1%	23.5%

**Tabela 3: Distribuição dos conjuntos de gestos ao longo da experiência**

## 5.1 Análise Quantitativa

A tabela 4 apresenta o tempo médio, em segundos, que os participantes demoraram para executar cada acção, tendo em consideração o conjunto de gestos.

	Sem Restrições	Moldura	Duplo Toque
Rectângulo	1.143	1.286	1.750
Elipse	0.696	0.750	1.167
Arrastar	1.108	1.054	1.107
<i>Scroll</i>	0.652	0.536	1.086
Seleccção	1.630	1.677	2.130
Copiar	0.111	0.403	1.002
Colar	0.211	0.208	0.498
<i>Undo</i>	0.341	0.550	0.806
<i>Redo</i>	0.350	0.375	0.605
Apagar	0.541	0.497	1.061

**Tabela 4: Tempo médio (em segundos) para cada acção em todos os conjuntos de gestos**

Para determinar a influência dos conjuntos de gestos na duração dos gestos realizados foram conduzidos testes ANOVA. A tabela 5 apresenta os resultados dos testes para todas as acções realizadas, menos as de desenho e de arrastar pois não implicam nenhum gesto de comando. O conjunto de gestos teve um efeito significativo nos gestos de *scroll*, copiar, *undo* e apagar com  $p < .001$ . Uma série de testes Tukey HSD foram realizados para comparar os gestos tendo em conta o conjunto pertencente. Estatisticamente concluímos que o gesto de *scroll* é significativamente mais lento no conjunto de gestos Duplo Toque do que no conjunto Moldura ( $HSD = 9.281, p < .01$ ) e no conjunto Sem Restrições ( $HSD = 7.924, p < .01$ ). O gesto de copiar também é significativamente mais lento no conjunto de gestos Duplo Toque do que no conjunto Moldura ( $HSD = 4.151, p < .01$ ) e no conjunto Sem Restrições ( $HSD = 6.663, p < .01$ ). O gesto de *undo* é significativamente mais rápido no conjunto Sem Restrições do que no conjunto Duplo Toque ( $HSD = 5.554, p < .01$ ). Finalmente, o gesto de apagar é significativamente mais lento no conjunto Duplo Toque do que no conjunto Moldura

(HSD = 6.262,  $p < .01$ ) e no conjunto Sem Restrições (HSD = 6.638,  $p < .01$ ).

	F-teste	$p$
<i>Scroll</i>	F(2.340) = 23.045	<b>4.1148E-10</b>
Seleção	F(2.120) = 4.248	0.016497981
Copiar	F(2.30) = 11.33	<b>0.000216102</b>
Colar	F(2.30) = 0.789	0.463676386
<i>Undo</i>	F(2.112) = 7.725	<b>0.000719604</b>
<i>Redo</i>	F(2.30) = 0.396	0.676793486
Apagar	F(2.120) = 13.424	<b>5.48109E-06</b>

**Tabela 5: Testes ANOVA para a duração dos gestos**

Em suma, os gestos de *scroll*, copiar e apagar são mais lentos no conjunto Duplo Toque comparativamente aos outros dois conjuntos. O gesto de *undo* é mais rápido no conjunto Sem Restrições quando comparado com o conjunto Duplo Toque. Estatisticamente não se encontraram mais relações significativas nos dados recolhidos.

Como referido anteriormente, a qualidade dos gestos realizados também foi medida. Para tal, um avaliador independente fez um reconhecimento visual de todos os gestos executados, cada gesto foi considerado de qualidade se este reconhecimento coincidia com a intenção do utilizador ao realizar o gesto. A tabela 6 apresenta para cada conjunto de gestos e cada acção a percentagem de gestos considerados de qualidade.

	Sem Restrições	Moldura	Duplo Toque
<i>Scroll</i>	92.1%	89.1%	78.6%
Copiar	94.4%	50.0%	60.0%
Colar	89.5%	60.0%	75.0%
<i>Undo</i>	72.7%	80.0%	54.8%
<i>Redo</i>	85.0%	87.5%	60.0%
Apagar	77.0%	76.7%	75.8%

**Tabela 6: Qualidade dos gestos para cada acção em cada conjunto de gestos**

Seguindo o que foi realizado para a duração dos gestos, executamos uma serie de testes ANOVA para determinar o efeito do conjunto de gestos na qualidade dos gestos realizados. A tabela 7 apresenta o resultado desses testes. Nesta tabela podemos observar que o conjunto de gestos tem um efeito estatisticamente significativo nos gestos de *scroll* e de copiar, com  $p < .05$ . É de observar que sendo estatisticamente significativo, este efeito é menos pesado do que o detectado no efeito sobre a duração dos gestos.

Mais uma vez, para comparar os gestos de *scroll* e de copiar realizamos testes Tukey HSD. Isto permitiu-nos concluir que o gesto de *scroll* no conjunto Sem Restrições apresenta um resultado estatístico de significante melhor qualidade do que no conjunto Duplo Toque, e que o gesto

de copiar apresenta uma melhor qualidade no conjunto Sem Restrições do que no conjunto Moldura.

	F-teste	$p$
<i>Scroll</i>	F(2.340) = 4.468	<b>0.012147763</b>
Copiar	F(2.30) = 4.574	<b>0.018461447</b>
Colar	F(2.30) = 1.748	0.191358791
<i>Undo</i>	F(2.112) = 2.807	0.064633036
<i>Redo</i>	F(2.30) = 0.921	0.408940323
Apagar	F(2.120) = 0.010	0.990287383

**Tabela 7: Testes ANOVA para a qualidade dos gestos**

Resumindo, o impacto do conjunto de gestos na qualidade dos gestos é menos profundo do que na média de tempo na execução do gesto.

## 5.2 Análise Qualitativa

Os dados qualitativos foram recolhidos através de dois questionários apresentados no fim de cada sessão experimental. No fim da primeira sessão foram recolhidos dados relativos à velocidade, ao conforto, à precisão e à satisfação referentes a cada gesto. Cada critério tinha uma escala de 1 a 5, sendo 5 a melhor pontuação. A tabela 8 apresenta os resultados médios dos 6 participantes.

	Sem Restrições	Moldura	Duplo Toque
Velocidade	4.33	3.00	2.67
Conforto	4.50	2.67	2.83
Precisão	4.17	2.83	3.33
Satisfação	4.33	2.50	2.67

**Tabela 8: Opinião dos participantes relativa aos três conjuntos de gestos**

Dado o pequeno número de participantes na experiência não foi feita análise estatística destes dados. Contudo, ao analisar a tabela podemos concluir que os participantes acharam que o conjunto Sem Restrições é mais confortável e rápido, tem grande precisão e de forma geral é mais satisfatório. Os resultados apresentam uma grande vantagem do conjunto Sem Restrições comparativamente com os outros dois conjuntos.

Ao comparar os outros dois conjuntos é possível verificar que não existe uma clara vantagem de nenhum deles sobre o outro, mesmo assim os participantes pareceram achar os gestos do conjunto Duplo Toque mais confortáveis de executar e mais precisos, enquanto que os gestos do conjunto Moldura mais rápidos de executar. Estes resultados estão de acordo com o esperado após uma simples análise a cada conjunto. Os gestos do conjunto Duplo Toque requerem a execução de um duplo toque no ecrã antes de executar o gesto de comando. Este requisito intuitivamente implica uma maior duração na execução do gesto, pelo menos em dispositivos onde a dimensão do ecrã não implique que, para aceder à moldura seja necessário um movimento muito longo. Esta análise é suporta-

da pelos dados da tabela 4, onde os gestos do conjunto Duplo Toque são os que requerem mais tempo de execução. No que respeita ao conforto e precisão, o requisito de executar o gesto a partir da moldura torna a sua execução mais desconfortável levando, logicamente, a que a precisão de execução diminua.

O questionário da segunda sessão foi focado em dois aspectos: qual o conjunto de gestos preferido e quão adequados são os gestos para cada acção. Para tal foi pedido aos participantes que escolhessem o seu conjunto de gestos preferido, aquele que consideraram mais fácil de executar, o mais eficiente e o mais fácil de memorizar. A resposta foi unânime em todas as questões: o conjunto Sem Restrições. Esta resposta está de acordo com a selecção feita pelos participantes quando a tarefa implicou uma livre escolha do conjunto, já que todos executaram os gestos do conjunto Sem Restrições.

A tabela 9 apresenta a satisfação dos participantes para cada gesto relativamente à acção desempenhada. A diferença entre os gestos residia somente na forma como era indicada a intenção de realizar um comando. A satisfação foi medida de 1 a 5, sendo 5 o melhor resultado.

	Satisfação
<i>Scroll</i>	4.00
Seleccção	2.83
Copiar	4.50
Colar	4.50
Apagar	3.50
<i>Undo</i>	5.00
<i>Redo</i>	5.00

**Tabela 9: Satisfação relativamente ao gesto associado a cada acção**

Os gestos de *scroll*, copiar, colar, *undo* e *redo* atingiram elevadas pontuações de satisfação, enquanto que os gestos de selecção e de apagar atingiram pontuações menores. Um dos participantes, sendo canhoto, comentou que teve dificuldades nestes dois gestos. Outro participante sugeriu que o gesto de apagar fosse multi-traço. O gesto de selecção criou alguns problemas por ser igual ao gesto de desenho de elipse variando somente na direcção. Esta diferença demonstrou não ser muito eficaz, justificando a baixa pontuação.

### 5.3 Observações e comentários

Comentários feitos pelos participantes durante conversas após os testes confirmaram aquilo que os questionários revelaram, o conjunto de gestos mais natural e menos intrusivo é o conjunto Sem Restrições. Esta é uma conclusão óbvia quando nos referimos a uma aplicação de desenho com um número limitado de gestos de desenho, mas ao serem adicionadas mais funcionalidades a aplicabilidade do conjunto Sem Restrições será menor. Se tomarmos como exemplo a possibilidade de desenhar linhas, da mesma forma que desenhámos rectângulos e

elipses, isto irá inviabilizar o gesto de *scroll*, uma vez que este também é uma linha. Um efeito parecido pôde ser observado durante a experiência. Sendo o gesto de desenhar uma elipse semelhante ao de selecção, por mais que uma vez os participantes executaram o gesto de selecção quando queriam desenhar uma elipse e vice-versa. Em aplicações de desenho reais será expectável que o número de gestos de desenho aumente e seja possível desenho livre, o que irá limitar acentuadamente o uso de gestos de comando sem restrições.

Uma observação relevante em sistemas gestuais é o impacto da mão dominante do utilizador. Sendo um dos participantes canhoto, foi possível observar, e o próprio comentou, que sentiu muitas dificuldades em distinguir o gesto de elipse do gesto de selecção, uma vez que variavam somente na direcção. Podemos também reparar que estes dois gestos geraram alguma confusão quando ocorria um erro de reconhecimento, levando os participantes a demorar mais tempo a pensar sobre o gesto correcto. Isto é um sinal inequívoco que a direcção do gesto não é suficiente para eliminar a ambiguidade entre gestos de desenho e de comando.

## 6. CONCLUSÕES E TRABALHO FUTURO

O foco principal deste trabalho foi como eliminar a ambiguidade existente entre gestos de comando e gestos de desenho em aplicações de desenho com interface gestual. Isto é particularmente importante em cenários em que não haja botões adicionais ou outra forma de “entrar” em modo comando. Propusemos duas dimensões, baseadas no início do gesto, para efectuar essa distinção. Uma dimensão era identificada pela localização do início do gesto, enquanto a outra por um duplo toque no ecrã antes do gesto de comando. A utilização destas dimensões permite que qualquer gesto seja utilizado como gesto de comando desde de que seja precedido por uma característica definida.

Para avaliar a nossa proposta foi conduzida uma experiência onde os dois conjuntos de gestos foram comparados entre eles, e também com uma situação ideal onde a intenção do utilizador era sempre reconhecida com 100% de eficácia, quer fosse desenhar ou executar um comando. Os conjuntos de gestos foram avaliados quanto à sua velocidade, conforto, precisão e satisfação. Tal como esperado, a situação ideal foi a escolha preferida dos nossos participantes assim como a mais efectiva.

De uma perspectiva pragmática de uma opção de implementação, a comparação entre os conjuntos Moldura e Duplo Toque revelou vantagens e desvantagens em ambos. A experiência revelou, na configuração em que foi realizada, que é mais rápido realizar gestos a partir da moldura do que com o duplo toque. No que respeita a conforto, precisão e satisfação os gestos precedidos de duplo toque demonstraram melhores resultados.

Futuramente iremos executar a experiência em diferentes dispositivos, desde telemóveis até ecrãs interactivos de grandes dimensões, tendo como objectivo estudar o

impacto da dimensão do ecrã na adequabilidade dos gestos.

Também será expectável que a ferramenta de desenho venha a ser complementada com mais algumas funcionalidades de desenho, aumentando assim o número de ambiguidades.

## 7. REFERÊNCIAS

- [Appert09] C. Appert and S. Zhai. Using strokes as command shortcuts: cognitive benefits and toolkit support. CHI '09, 2289-2298, 2009.
- [Balakrishnan04] R. Balakrishnan. "beating" fitts' law: virtual enhancements for pointing facilitation. *International Journal of Human-Computer Studies*, 61(6):857-874, 2004.
- [Bau08] O. Bau and W. E. Mackay. Octopocus: a dynamic guide for learning gesture-based command sets. UIST '08, 37-46, 2008.
- [Bragdon09] A. Bragdon, R. Zeleznik, B. Williamson, T. Miller, and J. J. LaViola, Jr. Gesturebar: improving the approachability of gesture-based interfaces. CHI '09, 2269-2278, 2009.
- [Cohen97] P. R. Cohen, M. Johnston, D. McGee, S. Oviatt, J. Pittman, I. Smith, L. Chen, and J. Clow. Quickset: multimodal interaction for distributed applications. MULTIMEDIA '97, 31-40, 1997.
- [Do-Lenh09] S. Do-Lenh, F. Kaplan, A. Sharma, and P. Dillenbourg. Multi-finger interactions with papers on augmented tabletops. TEI '09, 267-274, 2009.
- [Duarte09] C. Duarte and A. Neto. Gesture interaction in cooperation scenarios. CRIWG 2009, 190-205, 2009.
- [Fitzmaurice03] G. Fitzmaurice, A. Khan, R. Pieké, B. Buxton, and G. Kurtenbach. Tracking menus. UIST '03, 71-79, 2003.
- [Guimbretiére00] F. Guimbretiére and T. Winograd. Flowmenu: combining command, text, and data entry. UIST '00, 213-216, 2000.
- [Hinckley05] K. Hinckley, P. Baudisch, G. Ramos, and F. Guimbretiére. Design and analysis of delimiters for selection-action pen gesture phrases in scriboli. CHI '05, 2005.
- [Kurtenbach94] G. Kurtenbach and W. Buxton. User learning and performance with marking menus. In CHI '94, 258-264, 1994.
- [Landay96] J. A. Landay. Silk: sketching interfaces like crazy. CHI '96, 398-399, 1996.
- [Lane05] D. Lane, H. Napier, S. Peres, and A. Sandor. Hidden costs of graphical user interfaces: Failure to make the transition from menus and icon toolbars to keyboard shortcuts. *International Journal of Human-Computer Studies*, 18(1):133-144, 2005.
- [LaViola04] J.J. LaViola and R. C. Zeleznik. MathPad<sup>2</sup>: a system for the creation and exploration of mathematical sketches. *ACM SIGGRAPH 2004*, 432-440, 2004.
- [MacKenzie92] I. S. MacKenzie and W. Buxton. Extending fitts' law to two-dimensional tasks. CHI '92, 219-226, 1992.
- [Mankoff07] J. Mankoff, S. E. Hudson and G. D. Abowd. Interaction techniques for ambiguity resolution in recognition-based interfaces. *ACM SIGGRAPH 2007*, 2007.
- [Moyle03] M. Moyle and A. Cockburn. The design and evaluation of a flick gesture for 'back' and 'forward' in web browsers. AUIC '03, 39-46, 2003.
- [Mynatt99] E. D. Mynatt, T. Igarashi, W. K. Edwards and A. LaMarca. Flatland: new dimensions in office whiteboards. CHI '99, 346-353, 1999.
- [Neto09] A. Neto and C. Duarte. Comparing gestures and traditional interaction modalities on large displays. INTERACT 2009, 58-61, 2009.
- [Pedersen93] E. R. Pedersen, K. McCall, T. P. Moran and F. G. Halasz. Tivoli: an electronic whiteboard for informal workgroup meetings. CHI '93, 391-398, 1993.
- [Saund03] E. Saund and E. Lank. Stylus input and editing without prior selection of mode. UIST '03, 213-216, 2003.
- [Smith04] G. M. Smith and m. c. schraefel. The radial scroll tool: scrolling support for stylus- or touch-based document navigation. UIST '04, 53-56, 2004.
- [Wu03] M. Wu and R. Balakrishnan. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. UIST '03, 193-202, 2003.
- [Zeleznik06] R. Zeleznik and T. Miller. Fluid inking: augmenting the medium of free-form inking with gestures. GI'06, 155-162, 2006.

# Técnicas de Interacção para revisão de Cenários 3D

Bruno de Araújo

Diogo Mariano

Ricardo Jota

Alfredo Ferreira

Joaquim Jorge

Departamento de Engenharia Informática

INESC-ID/IST/Universidade Técnica de Lisboa

{brar,jota,afj,jaj}@vimmi.inesc-id.pt, diogo.mariano@ist.utl.pt

## Resumo

*A interacção com cenários tridimensionais usando ecrãs de grande dimensões continua um problema em aberto não apenas na academia, mas também na indústria. A nível de dispositivo continua-se à procura de uma alternativa ao rato e teclado de forma a tornar os sistemas de visualização mais interactivos e mais acessíveis aos utilizadores. As tecnologias de multi-toque e dispositivos sem fios vieram trazer novas alternativas, no entanto é necessário o redesenho das metáforas de interacção, dado não ser possível a simples adaptação das interfaces de Desktop para ecrãs de grandes dimensões. Neste trabalho propomos um conjunto de técnicas de interacção de forma a auxiliar a interacção com cenários tridimensionais com ecrãs de média e grande dimensão.*

## Palavras-Chave

*Interfaces Multi-toque, Dispositivos Multi-sensores, Ambientes Virtuais*

## 1 Introdução

A popularização da tecnologia de multi-toque e a evolução das interfaces lúdicas sem fios permite hoje em dia oferecer alternativas mais naturais para os utilizadores que lidam com conteúdo virtual a nível da prototipagem industrial, tal como a Arquitectura e a Indústria Automóvel. Dentro das áreas de concepção e desenho assistido por computador, os modelos virtuais são uma peça essencial permitindo não só auxiliar o processo de trabalho como oferecer uma alternativa aos protótipos físicos reduzindo o tempo de concepção, permitindo antecipar vários problemas de engenharia e construção. No entanto o processo de revisão desses protótipos limita-se muitas vezes ao uso dos programas de modelação utilizados na sua concepção, sendo baseados em interfaces pouco naturais e de longa aprendizagem, tornando o modelo virtual difícil de alterar e pouco acessível a utilizadores sem experiência nessas ferramentas de desenho assistido. Várias indústrias já recorrem a realidade virtual de forma a simular e experimentar protótipos, no entanto o seu uso foca principalmente na visualização sendo limitadas a nível da edição e manipulação do conteúdo virtual.

Neste trabalho pretende-se oferecer métodos alternativos de manipulação e edição de cenários tridimensionais de forma mais natural, tirando benefício de tecnologias que tornam a visualização mais interactiva e acessível ao utilizador. Por outro lado, pretende-se nesta fase inicial avaliar a adequação dos métodos apresentados de acordo com os requisitos obtidos pela análise de tarefas, antes de prosseguir para o desenvolvimento do protótipo final. Os ecrãs de grandes dimensões são cada vez mais

acessíveis, permitindo uma manipulação mais abrangente e colaboração entre vários utilizadores frente a uma superfície de visualização. De forma a ilustrar um conjunto de técnicas que melhoram as interacções com modelos virtuais, escolhemos dois cenários aplicados a duas indústrias diferentes. Por um lado, pretendemos oferecer um ambiente e métodos de interacção que auxiliem a Arquitectura, cuja principal média de comunicação são as plantas e as maquetas físicas, recorrendo a superfícies multi-toque. Por outro lado, pretendemos melhorar o processo de revisão a nível da indústria Automóvel, onde as ecrãs de grande dimensão já são comuns, embora actualmente a interacção se limite um operador que controla a visualização usando interfaces tradicionais *Desktop*. Nestes cenários pretendemos oferecer uma alternativa que tire partido da linguagem corporal, usando gestos com as mãos como alternativa à consola de controlo para realizar a revisão destes cenários. A abordagem proposta foi demonstrada e avaliada no âmbito do projecto Europeu MAXIMUS, permitindo a arquitectos e *designers* compor o ambiente virtual e alterar atributos dos objectos tridimensionais presentes na cena.

O artigo inicia-se com uma descrição do estado da arte relacionado com o uso de interfaces multi-toque e dispositivos 3D para interacção com cenários tridimensionais. Na secção 3 descrevemos os dois cenários realizados neste trabalho e como respondem às necessidades identificadas na revisão de protótipos virtuais em Arquitectura e na Indústria Automóvel. Na secção 4 descrevemos as técnicas de interacção que foram desenvolvidas e como os dispositivos de interacção foram utilizados em ambos os cenários. Finalmente apresentamos os resulta-

dos e análise dos primeiros testes com utilizadores, assim como as limitações e o trabalho futuro que vai ser realizado de forma a melhorar a interface actualmente proposta.

## 2 Trabalho Relacionado

Nos últimos anos vários trabalhos apresentaram métodos de interacção utilizando dispositivos de multi-toque para manipulação de conteúdo 2D, tais como imagens e vídeos. A maioria destes trabalhos utilizam um vocabulário de gestos semelhante baseado num ou em vários dedos de forma a realizar manipulações e rotações no plano de visualização. Wobbrock [Wobbrock 09] apresenta uma compilação de 23 gestos para manipular objectos 2D. No entanto, poucos trabalhos têm focado na adequação das interfaces multi-toque para interagir com cenários tridimensionais, e só recentemente este problema tem sido investigado [Hachet 08, Declec 09, Edelmann 09, Martinet 10]. Hachet [Hachet 08] apresenta um método que permite controlar os seis graus de liberdade da câmara virtual, utilizando exclusivamente coordenadas 2D. Nesta abordagem, o utilizador selecciona o ponto de interesse, desenhando um laço à volta do objecto e, automaticamente, o sistema apresenta uma esfera 3D que permite ao utilizador colocar a câmara em relação ao objecto. Durante a interacção com este ícone 3D, uma captura do ponto de vista é apresentada e actualizada interactivamente, auxiliando a colocação da câmara e permitindo navegar pelos objectos visíveis. Decle [Declec 09] apresenta uma solução alternativa utilizando o conceito de clique e arrastar em cima de um objecto 3D, permitindo explorá-lo. No entanto esta abordagem não permite examinar o objecto de perto, tornando difícil a visualização em cenários complexos compostos por vários modelos. Edelmann et al. [Edelmann 09] apresenta uma solução de interacção para cenários 3D utilizando dois movimentos, um para rodar a câmara (*pan* e *tilt*) e outro para se deslocar ou fazer zoom.

Recentemente Martinet [Martinet 10] apresenta uma solução para controlar o posicionamento de objectos, usando um dispositivo multi-toque combinando várias vistas ortogonais à semelhança dos sistemas de CAD. Esta solução apresenta um modo de manipulação preciso, no entanto obriga à utilização de demasiado espaço na área de visualização. De forma a ultrapassar os limites do plano multi-toque para o controlo dos seis graus de liberdade, vários trabalhos optaram por considerar o volume acima da superfícies da mesa para complementar a interacção. Através da colocação de uma câmara apontada para as mãos do utilizador no cimo da mesa, Hilliges [Hilliges 09] utiliza a estimação da distancia da mão à superfície de forma a agarrar e mover objectos, ultrapassando o limite da superfície. A abordagem é exemplificada utilizando a aplicação Google Earth, na qual uma mão permite mover o mapa e gestos com duas mãos permitem fazer zoom e controlar a elevação da câmara. Recorrendo a um sistema de *motion capture*, Marquardt [Marquardt 09] apresenta uma solução que não só permite interagir acima da mesa, como também combina o toque sobre a superfície,

permitindo efectuar movimentações e rotações de objectos. A maioria destes trabalhos apresentam uma interacção reduzida, focando normalmente apenas na navegação ou na manipulação de objectos. Por outro lado, nenhuma apresenta soluções dirigidas a Arquitectos, tal como sucede em cenários de interface tangíveis usando projecções de plantas [Ishii 02], ou cenários de realidade virtual usando mesas estereoscópica [Steinicke 06] ou óculos [Broll 03] para visualizar simulações de prédios virtuais.

No âmbito da interacção com grandes ecrãs, as soluções de multi-toque não são tão adequadas, em primeiro lugar devido ao problema tecnológico de escalabilidade para grandes dimensões, mas também devido à dimensão dos objectos, o que convida o utilizador a estar distante do ecrã e não em contacto directo. As primeiras abordagens optaram pela utilização de apontadores simples tais como lasers utilizando uma câmara [Lapointe 05] ou várias [Davis 02], permitindo aumentar a área de interacção. Esta solução é aplicada em vários trabalhos de larga escala apresentando o laser como uma boa alternativa ao rato [Oh 02], permitindo que vários utilizadores interajam com o ecrã [dA09a], adaptando o conceito da manipulação directa e interfaces *Desktop* [Olsen 01, Khan 04]. No entanto, o mapeamento dos 6 graus de liberdade para um dispositivo com 2 graus de liberdade continua limitado convidando ao uso de outros dispositivos. Hachet [Hachet 04] apresenta uma solução utilizando um prato circular de forma a mover e rodar objectos, mas obriga o utilizador a estar num local fixo em frente a ecrã. Reitmayr [Reitmayr 05] apresenta um dispositivo com fio de forma a controlar a rotação em ambiente de realidade aumentada, colocando um sensor com inércia dentro de uma bola. O aperto da bola permite interagir com menus circulares. Baudisch [Baudisch 06] apresenta um dispositivo sem fios semelhante a uma pequena esponja como simples alternativa ao rato. Ainda assim, torna-se necessário um dispositivo mais completo para permitir a correcta manipulação de objectos virtuais controlando os seis graus de liberdade, motivo pelo qual optámos por propor recentemente a SqueazyBall [dA09b] para interacção com ecrãs de larga escala.

## 3 Cenários e Dispositivos de Interacção

Este trabalho enquadra-se no Projecto Europeu Maximus que pretende explorar novos paradigmas de interacção de forma a tornar mais natural a interacção com ecrãs de média e grandes dimensões, com o objectivo de suportar a revisão de protótipos virtuais de Arquitectura e da Indústria Automóvel.

Foram desenhados dois cenários de interacção analisando a forma como os arquitectos e construtores de automóveis utilizam os modelos virtuais e físicos no seu processo de trabalho. Para tal analisou-se como os utilizadores usam actualmente as ferramentas existentes, tais como modeladores de CAD para compor cenários tridimensionais e configurar os vários graus de realismo da visualização. Esta análise combinada com o estudo do uso de modelos físicos permitiu identificar funcionalidades de



Figura 1: Ambiente de trabalho de um Arquitecto e sala de projecção de Design Automovel

manipulação de objectos, materiais, luzes e câmaras, assim como um conjunto de gestos e pistas utilizadas pelos *designers* durante o processo de revisão de um modelo virtual.

No âmbito da Arquitectura, a revisão de modelos envolve vários documentos, tais como plantas de edifícios, modelos virtuais renderizados com realismo e maquetes físicas, permitindo a vários utilizadores discutir as opções de desenho. No processo conceptual de desenho, o uso de modelos tridimensionais é limitado à produção de imagens e visitas virtuais. Quanto às discussões entre arquitectos, ocorrem normalmente em frente a plantas em tamanho A0 com um conjunto de imagens ou documentos apresentando informação adicional sobre o projecto (Figura 1 esquerda). Em relação aos modelos físicos, são custosos e estáticos e as maquetes são normalmente utilizadas para apresentação do projecto, podendo também servir para experimentar um conceito, tendo pouco peso no processo de desenho. As ferramentas de CAD propõem poucas alternativas para auxiliar a revisão, no entanto são necessárias para transpor as alterações no modelo virtual.

Ao contrário da Arquitectura, a Indústria Automóvel, tal como a Aeronáutica, foram pioneiras na utilização de ambientes virtuais no seu processo de desenho, assim como no uso de ecrãs de grandes dimensões. Hoje em dia é comum o uso de *showrooms* (Figura 1 direita) para a visualização de modelos virtuais de um protótipo utilizando ecrã de larga escala, o que permite ver automóveis à escala de um para um com grande realismo. A utilização de modelos virtuais permite avaliar o desenho e a estética reduzindo a necessidade de recorrer a modelos físicos de alto custo tanto a nível económico como temporal. No entanto durante as sessões de revisão, as interacções para navegação e alteração de desenho são

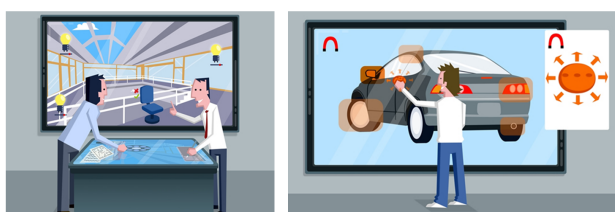


Figura 2: Cenários de interacção com ecrã de média e grande dimensão

limitadas e normalmente restritas a um operador que controla a visualização seguindo as instruções dos *designers*. As próximas secções apresentam os dois cenários desenvolvidos assim como os dispositivos utilizados, propondo uma solução alternativa a revisão de protótipos virtuais.

### 3.1 Cenário para Arquitectura

O cenário dos Arquitectos centra-se no uso de dois dispositivos: uma mesa multi-toque para interacção e visualização de plantas e um ecrã de projecção mostrando uma renderização de alta qualidade (Figura 2 esquerda). A mesa multi-toque foi construída recorrendo a três espelhos para encurtar a projecção e oferecer um tamanho A0 de visualização, respondendo ao requisito dos arquitectos que desejavam um tamanho semelhante ao das suas plantas.

O suporte multi-toque foi inserido seguindo a abordagem *Lazer Light Plan* apresentada pelo grupo NuiGroup<sup>1</sup>, onde são usados lasers de luz infra-vermelha dispostos sobre a superfície, que com a ajuda de um gerador de linha criam um plano de luz infra-vermelha sobre a mesma. Quando o utilizador toca na superfície, esse plano é perturbado e uma câmara localizada no interior da mesa capta essa alteração. Utilizando a ferramenta TBeta do NuiGroup são determinados as posições onde ocorrem os toques. Com esta abordagem é possível que vários utilizadores interajam em simultâneo sem existirem limitações quanto ao número de toques suportados, quando comparado a outros suporte físicos de multi-toque. Quanto ao ecrã de projecção adicional permite visualizar renderizações realistas do mundo virtual e apresentar visitas virtuais para complementar a interacção com a planta. A Figura 3 mostra a realização do cenário de interacção para auxiliar os Arquitectos nas tarefas de revisão. Com este cenário pretende-se que seja possível aos arquitectos manipular as plantas na mesa multi-toque, num ambiente colaborativo, e que as alterações que eles realizam sejam visíveis no ecrã de larga escala sem necessitar recorrer a outro sistema.

<sup>1</sup>[http://wiki.nuigroup.com/Laser\\_Light\\_Plane\\_Illumination\\_\(LLP\)](http://wiki.nuigroup.com/Laser_Light_Plane_Illumination_(LLP))



Figura 3: Dois utilizadores visualizando uma planta na mesa multi-toque e um vista de uma divisão de uma casa





Figura 4: Utilizador inspecionando o modelo de um automóvel usando a SqueezyBall

### 3.2 Cenário para Indústria Automóvel

Para a indústria Automóvel, o cenário desenhado foca no uso de um dispositivo multi-sensor sem fios chamado SqueezyBall [dA09b] para interagir com um ecrã de larga escala (Figura 2 direita). Este dispositivo pretende ser de fácil utilização complementando os *showrooms* existentes. O formato de uma bola, permite ao utilizador abstrair-se do estado do dispositivo e largá-lo ou partilhá-lo sem preocupação. Além disso, a bola incentiva o uso de gestos naturais tais como a selecção de objectos, o controlo de orientações e de intensidades das luzes virtuais rodando a bola. Para activar elementos e opções optámos pelo gesto de apertar utilizando um sensor de pressão dentro da bola. Por outro lado, o uso de acelerómetros e um giroscópio permite reconhecer gestos efectuados pelo utilizador assim como a orientação da própria bola. Utilizando um sistema de *motion capture* é possível determinar qual a posição do dispositivo enquanto o utilizador agarra a bola.

A SqueezyBall foi desenvolvida recorrendo ao uso de uma plataforma multi-sensores da empresa Plux<sup>2</sup>, permitindo incluir vários sensores apresentados em [dA09b]. Desta forma foi possível responder aos requisitos dos utilizadores que desejavam ter a possibilidade de agarrar objectos apertando o dispositivo, definir rotações rodando o dispositivo, movimentar objectos virtuais e realizar gestos movimentando a bola no espaço ou accionar certas opções do sistema utilizando botões. A SqueezyBall utiliza o protocolo de comunicação *Bluetooth* e inclui uma bateria para alimentar os vários sensores localizados no interior da bola. Desta forma é possível uma interacção no espaço sem restrições de fios, estando todos os sensores escondidos no interior da bola. A detecção da posição é efectuada de forma transparente para o utilizador, dado a bola ser revestida de material reflector ao infra-vermelho. Utilizando um sistema de câmaras é possível de forma não intrusiva seguir a bola sendo esta vista como uma marca ou um conjunto de marcas tridimensionais agregadas, de forma a calcular a sua posição no espaço. Sabendo a posição da bola é possível saber onde a mão do utilizador se encontra em relação ao ecrã de grandes dimensões e in-

<sup>2</sup><http://www.plux.info>

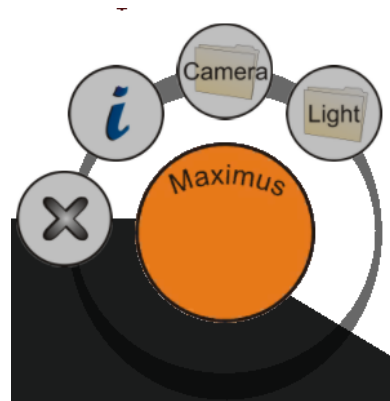


Figura 5: Menu principal da Aplicação em formato de disco.

teragir com este sem necessitar outro dispositivo. A bola pode ser partilhada por vários utilizadores durante uma sessão de revisão de um protótipo virtual sem necessidade de intervenção de um operador para auxiliar a visualização exemplificada na Figura 4.

## 4 Técnicas de Interação

Os dois cenários apresentados na secção anterior foram concretizados utilizando o mesmo protótipo, de forma a permitir ao utilizador manipular entidades de um cenário tridimensional e configurar atributos da visualização tais como materiais de objectos, características das luzes e posicionamento de câmaras pelo cenário virtual. De seguida apresentamos a interface de utilizador e as técnicas de interacção para cada cenário. No primeiro cenário focámos a interface para a mesa multi-toque e de seguida introduzimos as técnicas de interacção com um ecrã de larga escala utilizando a SqueezyBall.

### 4.1 Interface Utilizador

A Interface Utilizador da aplicação Maximus é constituída por dois tipos de menus, que permitem ao utilizador interagir com a aplicação e manipular propriedades dos objectos, alterando o estado global do sistema. O menu principal da aplicação encontra-se sempre visível e menus contextuais podem ser abertos consoante o objecto, de forma a configurar seus atributos. Optámos por utilizar menus circulares semi-transparentes apresentando um número limitado de opções e recorrendo no máximo a uma hierarquia de dois níveis. O segundo nível pode ser activado seleccionado a respectiva opção no primeiro nível. A semi-transparência permite ver os objectos da cena localizados por trás do menu e a disposição circular permite reduzir a distância entre as opções.

O menu principal, apresentado na Figura 5, permite aceder a qualquer luz ou câmara mesmo quando esta não se encontra visível. A parte central circular permite seleccionar o menu podendo ser movimentado, escalado ou rodado caso o utilizador o deseje. O menu tem um aspecto semelhante a um disco, permitindo no futuro incluir comportamentos físico para uma interacção mais natural sobre a mesa multi-toque.



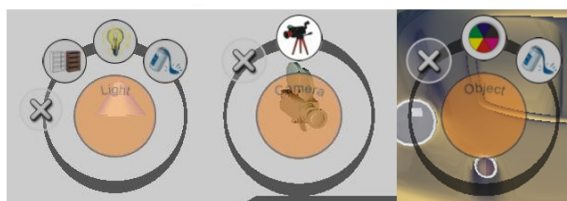


Figura 6: Menus contextuais da aplicação: para a configuração de luzes, activação de câmara e configuração de objectos

Relativo aos menus contextuais, estes apresentam o mesmo visual que menu principal, no entanto só apresentam opções relacionadas com a selecção por parte do utilizador de um dado tipo de objecto. A Figura 6 apresenta os três tipos de menus contextuais existentes. No menu contextual de luzes é possível activar ou desactivar, mudar a intensidade ou a cor da luz. O menu de câmara permite activar uma dada câmara. Quanto ao menu dos objectos, este permite esconder, apagar ou alterar o material de um dado modelo através de uma paleta de cores.

## 4.2 Interface Multi-toque

A interacção no cenário dos arquitectos é feita com a ajuda da superfície multi-toque. A ideia base centra-se na possibilidade de efectuar um conjunto de diferentes gestos sobre qualquer entidade manipulável existente na planta de forma a movimentar, rodar ou escalá-la. O dicionário de gestos foi definido baseando-se na linguagem comum disponível em vários ecrãs multi-toque ou dispositivos móveis, simplificando a criação de modelos mentais de interacção por parte dos utilizadores.

Dado que os menus utilizam a mesma linguagem gestual, é possível a vários utilizadores estarem dispostos em torno da mesa, a utilizar a aplicação movendo, escalando e orientando os menus de acordo com a sua posição. A interface gestual de manipulação de objectos é composta por cinco gestos ilustrados na Figura 7. Utilizando um dedo é possível transladar objectos por manipulação directa. Usando dois dedos é possível rodar um objecto no eixo perpendicular á superfície, ou nos outros eixos utilizando dois dedos próximos semelhante a um gesto de *scroll*. Afastando ou aproximando dois dedos é possível escalar o objecto de forma uniforme. Finalmente, os menus contextuais são instanciados mantendo o dedo sobre um objecto durante dois segundos. Este princípio também foi utilizado como primeira abordagem para a selecção de opções de menu. O reconhecimento é efectuado aglomerando os dedos por proximidade e usando as propriedades geométricas do gesto comparado ao estado anterior.

## 4.3 Interface SqueezyBall

Para a interacção com ecrãs de largas dimensões, o volume onde se situa o utilizador é obtido usando um sistema de *motion capture* para capturar a posição tridimensional da SqueezyBall em relação ao ecrã. A SqueezyBall age como um cursor visível para o utilizador projectando a posição

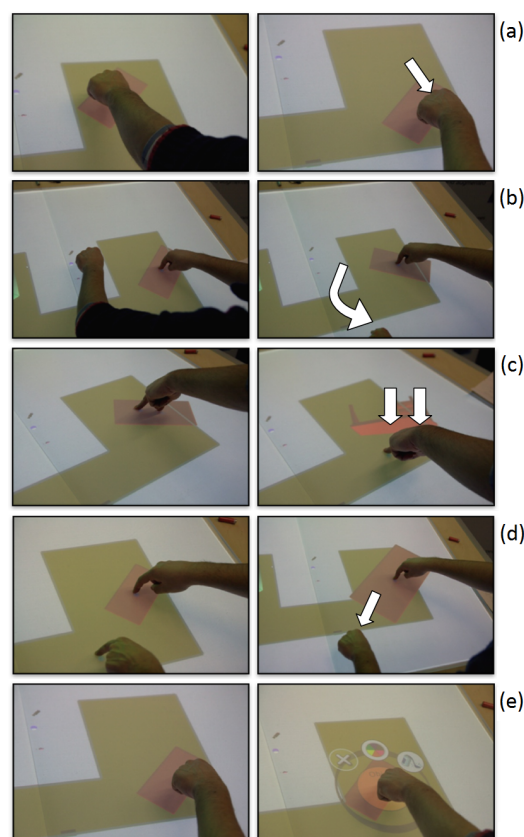


Figura 7: Linguagem gestual utilizada pela interface Multi-toque: a) movimentação de objecto, b) rotação no eixo perpendicular a mesa, c) rotação nos outros eixos, d) escalonamento e e) abertura de menu contextual

da bola de forma ortogonal na área de visualização. O cursor é utilizado de forma a informar qual o estado do dispositivo permitindo distinguir entre movimentações e mudanças de orientação de objectos. O utilizador pode manipular qualquer entidade apresentada no ecrã, podendo corresponder a um objecto da cena tal como um carro no cenário da Indústria Automóvel, uma luz ou uma câmara do mundo virtual. Por outro lado o dispositivo pode ser utilizado para interagir com menus e aceder a funcionalidades adicionais de cada entidade.

Para manipular um dado objecto tridimensional, o utilizador simplesmente coloca o cursor sobre o objecto pretendido deslocando a SqueezyBall livremente no ar e selecciona-o apertando a bola. O sensor de pressão da bola permite detectar quando é que um objecto está a ser manipulado. Para deslocar um objecto basta seleccionar apertando a bola, movimentar e relaxar a bola, colocando o objecto na nova posição. Este processo é semelhante à interacção física com objectos reais seguindo a analogia conhecida do *Drag and Drop*. Quando movimentos são efectuados em profundidade, é possível deslocar o objecto nessa direcção. Esta medida é calculada de forma relativa referenciando a profundidade de origem do objecto. Este tipo de deslocamento na cena foi mapeado num espaço de um metro permitindo ao utilizador facilmente

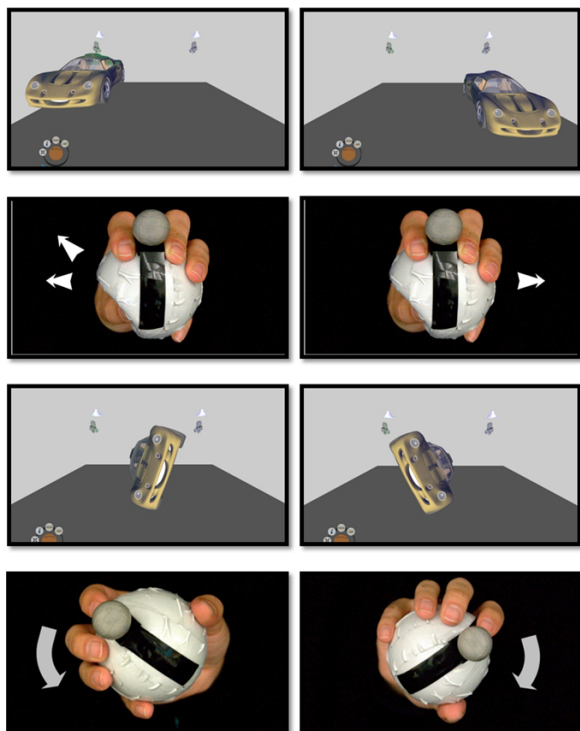


Figura 8: Exemplos de manipulações utilizando a SquezyBall: Movimentação e Rotação de Objectos

alterar a profundidade do objecto a partir da sua posição. A interacção com menus efectua-se da mesma forma, colocando o cursor através da posição da bola. Para seleccionar uma opção, basta colocar o cursor sobre a opção e pressionar o botão da SquezyBall. Esta interacção é válida para o menu global e para os menus contextuais correspondentes a cada objecto. Para abrir o menu contextual de um objecto, basta colocar a SquezyBall sobre este e pressionar o botão duas vezes.

Para rodar um objecto, o utilizador selecciona-o apertando a bola duas vezes seguidas. O cursor informa o utilizador que a bola se encontra em modo de rotação, permitindo-lhe rodar o objecto rodando a bola. A rotação é calculada utilizando os acelerómetros e o giroscópio da SquezyBall e é feita de forma directa mapeando a orientação relativa da bola na orientação do objecto. A rotação é finalizada largando a bola. Através da utilização do apertar (Figura 8), da posição e orientação da SquezyBall assim como do botão localizado na bola, é possível sem necessidade de outro dispositivo aceder a todas as funcionalidades da aplicação e controlar os vários objectos do cenário tridimensional.

## 5 Avaliação Preliminar

De modo a validar a abordagem seguida, foi realizada uma avaliação experimental do protótipo envolvendo utilizadores reais. Nesta avaliação pretendia-se verificar também se os requisitos da análise de tarefas estavam a ser cumpridos de acordo com as expectativas dos utilizadores. Tendo em consideração que o protótipo avaliado consistia numa versão preliminar, os principais objectivos dos testes

com utilizadores foram a apresentação a utilizadores reais da versão actual do protótipo e a recolha de informação que permitisse identificar o caminho a seguir na fase seguinte de desenvolvimento. De realçar que o protótipo final, desenvolvido a partir da versão avaliada de acordo com a informação reunida na avaliação preliminar, será sujeito a uma avaliação final, planeada para o final do ano com mais arquitectos e designers de automóveis.

### 5.1 Testes com Utilizadores

Nesta avaliação do protótipo foram envolvidos sete utilizadores, dos quais seis eram do sexo masculino. Este grupo incluía um designer de automóveis e dois arquitectos avaliando nesta fase preliminar a adequação dos métodos apresentados em relação aos requisitos dos utilizadores. Os utilizadores tinham entre 25 a 34 anos e todos possuíam experiência com ferramentas de modelação. No entanto, a grande maioria dos utilizadores (seis) tinham uma fraca experiência com ecrã multi-toque. Os utilizadores afirmaram, ter uma média experiência com dispositivos apontadores tais como o Wiimote. Para a realização dos testes foram preparados dois cenários distintos. Um cenário dedicado ao contexto dos designers de automóveis e outro ao dos arquitectos. Estes cenários foram criados em salas separadas. A sala dedicada ao cenário automóvel continha um projector HD e um sistema de tracking, de forma a seguir a Squezy Ball. Na outra sala simulava-se o cenário de arquitectura utilizando a mesa multi-toque e um ecrã vertical retro-projectado para uma vista adicional da cena. Parte da realização da avaliação consiste em convidar o utilizador a realizar uma sequência de tarefas varrendo as várias funcionalidades e as metáforas de interacção propostas pelo protótipo.

O teste iniciou-se com a apresentação do protótipo numa sessão em que todos os utilizadores participaram. Em simultâneo, foi distribuído um manual da aplicação, onde eram apresentados os paradigmas de interacção e descritas de forma sucinta as funcionalidades oferecidas pela aplicação. Cada utilizador teve então oportunidade de se ambientar com o protótipo, experimentando-o livremente. De seguida foram realizadas as sessões de testes formais do protótipo. Estas consistiam na realização de uma sequência de dois conjuntos de tarefas de dificuldade variável. Um conjunto de oito tarefas no contexto do desenho automóvel e um conjunto de nove tarefas no cenário de arquitectura. As tarefas consistiam na movimentação e orientação de objectos do cenário assim como da alteração de atributos tais como a intensidade das luzes ou os materiais dos objectos. Por outro lado, eram pedidas alterações da vista de forma a ver detalhes específicos do modelo tridimensional tal como acontece em sessões de revisão de projecto. Durante essas sessões, os observadores mediram o tempo de execução e erros efectuados pelos utilizadores. Adicionalmente, as várias sessões foram gravadas em vídeo. Na sequência de cada sessão, os utilizadores foram convidados a responder a um inquérito de forma a recolher mais informação sobre a interface.

## 5.2 Inquéritos e Entrevistas

Neste inquérito foram utilizados questionários de usabilidade seguindo a norma ISO 9241, parte 10. Esta norma baseia-se na avaliação de sete factores: adequação do sistema à tarefa, auto-descritividade da interface, controlabilidade do sistema, conformidade a expectativa do utilizador, tolerância a erros, adequação as necessidades do utilizador e facilidade de aprendizagem. A segunda parte dos questionários tinha por objectivo quantificar os benefícios do protótipo e a sua facilidade a nível criativo. Para complementar a realização das tarefas e a resposta aos questionários, foi realizada uma sessão de entrevistas com os utilizadores, depois dos testes. Desta forma conseguimos recolher comentários e sugestões que poderiam não ter sido transmitidas durante os inquéritos. Conseguimos assim, com base nas respostas aos questionários recolher informação sobre atributos relacionado com hardware, software e o ambiente apresentado que contribuem para a usabilidade e os seus princípios ergonómicos.

## 5.3 Análise de Resultados

No cenário da mesa, cada utilizador executou sete tarefas com dificuldade variável. A realização das tarefas mostrou que uma das limitações é a falta de re-alimentação dos gestos efectuados por parte da interface actual, com excepção da não alteração da posição e orientação dos modelos manipulados, não houve qualquer outra re-alimentação, em particular quando um toque era reconhecido pelo sistema. Este ponto sugere que não só deve ser oferecida informação sobre a presença de toque na mesa, como também deve ser dada por parte do sistema informação sobre o gesto que está a ser reconhecido mesmo durante uma manipulação directa. Isso aumentaria a aprendizagem do utilizador permitindo maior controlo das manipulações, evitando, por outro lado, movimentações involuntárias. Verificamos que as operações relacionadas com a manipulação de pequenos objectos demoram mais tempo devido a dificuldade de selecção obrigando a recorrer ao zoom para certos objectos. Por um lado a manipulação e orientação de objectos revelou-se fácil, mas o controlo de câmara não pareceu suficientemente natural, sendo necessária outra analogia tal como a manipulação do ponto de foco em vez da rotação livre da câmara. Por outro lado a não restrição dos graus de liberdades na rotação, levou muitas vezes a orientações incoerentes para o utilizador, sendo necessário um melhor controlo dos graus de liberdade da rotação. Relativo a interacção com os menus, a sua usabilidade deve ser melhorada a nível da selecção de forma a evitar o uso de tempo de espera para a activação de opções. O número reduzido de opções e níveis de menu tornou-se benéfico assim como a sua fácil movimentação e redimensionamento quando necessário. Em relação aos objectos, a manipulação de menus pareceu mais natural por ser uma entidade de duas dimensões.

Quanto à utilização da SqueezyBall, os utilizadores encontraram dificuldades na execução da acção de double click do botão para a abertura dos menus contextuais e do apertar

duplo para activar a rotação dos modelos. As acções duplas (clique do botão ou apertar) requerem que sejam executadas num espaço de tempo demasiado restritivo para ser considerado como uma acção atómica dupla. A sensibilidade do reconhecimento de acções duplas deve ser configurada e adaptada pelo utilizador. Por outro lado, quando o utilizador executa um duplo apertar da bola, a Squeezy Ball é lenta a recuperar o seu estado original devido ao fluxo de ar no seu interior, o que provoca tempos mortos na sua utilização caso o utilizador aperte demais a bola. Por esse motivo as acções duplas tornam-se difíceis e a solução actual de escape de ar deve ser revista de forma a oferecer um apertar mais natural ao utilizador. Quanto às funcionalidades oferecidas pelo sistema, para além dos problemas na manipulação de objectos, a manipulação das câmaras revelou-se pouco natural. De facto, metade dos utilizadores movimentavam o carro em vez da câmara. A representação de vistas adicionais das câmaras, não permitiu evitar esse problema sendo necessário melhorar a funcionalidade de navegação de forma a evitar movimentações de objectos indesejadas. Quanto à rotação de elementos geométricos verificamos que nem sempre o mapeamento directo é natural para o utilizador. Alguns utilizadores tentavam efectuar um gesto de rotação com a mão deslocando o cursor em vez de rodar a própria bola. Este problema alerta para a necessidade de uma representação gráfica da rotação que deve ser apresentado de forma a facilitar a sua percepção e auxiliar o utilizador a recuperar de uma rotação indesejada.

## 6 Conclusão

Este trabalho propôs duas metáforas de interacção para a revisão de cenário tridimensionais utilizando ecrãs de média e grandes dimensões. A primeira, direccionada à Arquitectura, mostrou a viabilidade e a mais valia para os utilizadores de uma interface multi-toque combinada com uma projecção adicional para interagir com plantas, auxiliando a revisão de projectos. No segundo mostrámos a vantagem do uso de dispositivo multi-sensores de forma a interagir com um ecrã de grandes dimensões. Como trabalho futuro pretendemos ultrapassar algumas limitações relacionadas com a falta de re-alimentação das interfaces gestuais e permitir mais funcionalidades tais como a visualização de secções e definição de caminhos de câmara. Por outro lado pretendemos incluir suporte para anotações. Igualmente desejamos considerar restrições de graus de liberdade tanto a nível do multi-toque como da SqueezyBall, dado que a interacção livre nem sempre oferece a precisão necessária. Num futuro próximo planeamos realizar novos testes de forma a avaliar o desempenho destas abordagens comparado aos métodos tradicionais.

## 7 AGRADECIMENTOS

Este trabalho foi parcialmente patrocinado pela Fundação para a Ciência e Tecnologia (SFRH/BD/31020/2006 e 17574/2004) e pelo projecto MAXIMUS (IST-2007-1-217039) parte do FP7 da Comunidade Europeia.

## Referências

- [Baudisch 06] P. Baudisch, M. Sinclair, e A. Wilson. SOAP: a pointing device that works in mid-air. Em *19th ACM Symp. on User Interface Software and Technology*, p. 43–46, New York, NY, USA, 2006. ACM.
- [Broll 03] W. Broll, M. Stoerring, e C. Mottran. The augmented round table - a new interface to urban planning and architectural design. Em *INTERACT '03: IFIP TC13 International Conference on Human-Computer Interaction*. IOS Press, 2003.
- [dA09a] B. de Araújo, R. Jota, L. Bruno, J. A. Madeiras Pereira, e J. A. Jorge. Immiview - a multi-user solution for design review in real-time. *Journal of Real-Time Image Processing: Special Issue on Improving Display and Rendering Technology*, 2009.
- [dA09b] B. de Araújo, R. Jota, J. Fernandes, A. Ferreira, e J. Madeiras Pereira. Dispositivo de interacção para ecrã de grandes dimensões. Em *Poster no 17º Encontro Português de Computação Gráfica*, 2009.
- [Davis 02] J. Davis e X. Chen. Lumipoint: Multi-user laser-based interaction on large tiled displays. *Displays*, 2002.
- [Declec 09] F. Declec, M. Hachet, e P. Guitton. Technote: Scruticam: Camera manipulation technique for 3d objects inspection. *3D User Interfaces*, 0:19–22, 2009.
- [Edelmann 09] J. Edelmann, A. Schilling, e S. Fleck. The dabr: A multitouch system for intuitive 3d scene navigation. Em *3DTV09*, p. 1–4, 2009.
- [Hachet 04] M. Hachet, P. Guitton, P. Reuter, e F. Tyndiuk. The CAT for efficient 2d and 3d interaction as an alternative to mouse adaptations. *ACM Trans. Graph.*, 23(3):731–731, 2004.
- [Hachet 08] M. Hachet, F. Decle, S. Knodel, e P. Guitton. Navidget for easy 3d camera positioning from 2d inputs. Em *3DUI '08: IEEE Symposium on 3D User Interfaces*, p. 83–89, Washington, DC, USA, 2008. IEEE Computer Society.
- [Hilliges 09] O. Hilliges, S. Izadi, A. D. Wilson, S. Hodges, A. Garcia-Mendoza, e A. Butz. Interactions in the air: adding further depth to interactive tabletops. Em *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*, p. 139–148, New York, NY, USA, 2009. ACM.
- [Ishii 02] H. Ishii, E. Ben-Joseph, J. Underkoffler, L. Yeung, D. Chak, Z. Kanji, e B. Piper. Augmented urban planning workbench: Overlaying drawings, physical models and digital simulation. *Mixed and Augmented Reality, IEEE / ACM International Symposium on*, 0:203, 2002.
- [Khan 04] A. Khan, G. Fitzmaurice, D. Almeida, N. Burtnyk, e G. Kurtenbach. A remote control interface for large displays. Em *UIST '04: ACM symposium on User interface software and technology*, p. 127–136, New York, NY, USA, 2004. ACM.
- [Lapointe 05] J. Lapointe e G. Godin. On-screen laser spot detection for large display interaction. *HAVE*, 2005.
- [Marquardt 09] N. Marquardt, R. Jota, S. Greenberg, e J. Jorge. The continuous interaction space: Integrating gestures above a surface with direct touch. Relatório Técnico 2009-925-04, Department of Computer Science, University of Calgary, Canada, April 2009.
- [Martinet 10] A. Martinet, G. Casiez, e L. Grisoni. The design and evaluation of 3d positioning techniques for multi-touch displays. Em *3D User Interfaces (3DUI), 2010 IEEE Symposium on*, p. 115 –118, 2010.
- [Oh 02] J. Oh e W. Stuerzlinger. Laser pointers as collaborative pointing devices. *CHI*, 2002.
- [Olsen 01] D. R. Olsen e T. Nielsen. Laser pointer interaction. *SIGCHI*, 2001.
- [Reitmayr 05] G. Reitmayr, C. Chiu, A. Kusternig, M. Kusternig, e H. Witzmann. iorb - unifying command and 3d input for mobile augmented reality. Em *IEEE VR 2005 Workshop on New Directions in 3D User Interfaces*, 2005.
- [Steinicke 06] F. Steinicke, T. Ropinski, K. H. Hinrichs, e J. Mensmann. Urban city planning in semi-immersive virtual reality. Em *Proceedings of the International Conference on Computer Graphics Theory and Applications (GRAPP2006)*, p. 192–199, Setúbal, 2006.
- [Wobbrock 09] J. Wobbrock, M. R. Morris, e Andrew D. Wilson. User-defined gestures for surface computing. Em *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, p. 1083–1092, New York, NY, USA, 2009. ACM.

# Interacção Multi-toque no contexto do DJing

Pedro A. Lopes J. A. Madeiras Pereira Alfredo Ferreira

Departamento Engenharia Informática

INESC-ID/IST/Universidade Técnica de Lisboa

R. Alves Redol, 9, 1000-029 Lisboa, Portugal

pedro.lopes@ist.utl.pt, jap@inesc-id.pt, alfredo.ferreira@inesc-id.pt

## Resumo

Nos últimos anos têm-se procurado novas formas de interacção para os Disc-Jockeys (DJs), no entanto poucas soluções conseguem capturar a sinergia entre os benefícios da interacção multi-toque e a virtualização do ambiente de manipulação áudio, evitando o input através de dispositivos tradicionais, como o rato ou o teclado. Centrado no envolvimento de um grupo de DJs profissionais, neste artigo propomos um sistema modular, caracterizado pela interacção natural, consistente com os gestos tradicionalmente usados e sem necessidade de objectos tangíveis adicionais. O sistema proposto foca os principais desafios de interacção em aplicações de DJing virtual, usando superfícies multi-toque.

## Palavras-Chave

Superfície Multi-toque, DJ, Desafios Interacção, Interface, Aplicação Modular, Concepção Centrada no Utilizador

## 1. Introdução

A actividade do DJ contemporâneo ultrapassou há muito a simples escolha de discos em ambientes festivos e, actualmente, o DJ é possuidor de conhecimentos técnicos extensos e capaz de executar tarefas complexas na interacção com o hardware que utiliza. Durante três décadas este equipamento (denominado *DJ Setup*) manteve-se quase inalterável, sendo tipicamente composto por fontes sonoras (gira-discos, leitores de CD), uma mesa de mistura (dispositivo capaz de misturar pelo menos dois sinais áudio) e discos (registos no formato físico: vinil ou CD) [7].

Com o virar do milénio, operou-se uma revolução na digitalização do *Setup Tradicional* dos DJs, fruto da massificação dos computadores portáteis. No entanto, as aplicações comuns de DJing virtual possuem um modelo de interacção não-natural, que recorre tipicamente a dispositivos de entrada tradicionais (rato e teclado) ou controladores hardware MIDI. Assim sendo, torna-se difícil a interacção com os utilizadores, consequência da disparidade entre os comandos virtuais e as acções já aprendidas do *Setup Tradicional*. Observa-se nestas aplicações um aumento da curva de aprendizagem limitando-se a actividade, tipicamente bimanual, a um fluxo de controlo único.

Ainda na primeira metade da década transacta a indústria concebeu um novo tipo de sistema, capaz de minimizar os problemas de interacção do *Setup Virtual*; tal sistema é composto pela união dos anteriores, como está patente na Figura 1, e denomina-se *Setup Híbrido*. Trata-se de um sistema digital de controlo de uma aplicação de DJing virtual através dos componentes típicos: gira-discos ou leitores

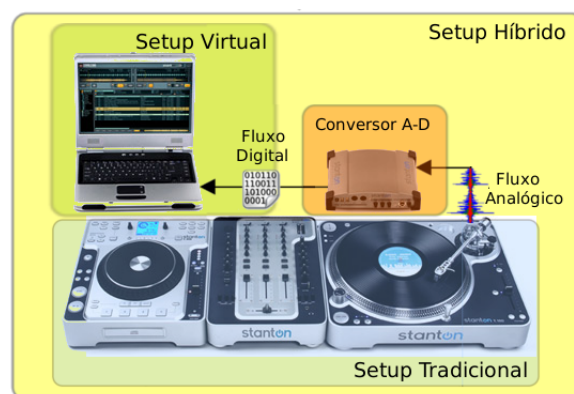


Figura 1. Setup Tradicional, Virtual e Híbrido.

de CD. Neste *Setup* o utilizador pode controlar o *playout* do ficheiro digital, manipulando directamente o disco de vinil que contém informação codificada de *tracking*, o que permite ao software determinar posição do ficheiro áudio. Embora os problemas inerentes à manipulação não-natural do *Setup Virtual* sejam resolvidos, tais soluções híbridas voltaram a despoletar a necessidade de recorrer aos limitados equipamentos analógicos tradicionais com elevados custos de manutenção, aquisição e transporte associados.

Para minimizar estas desvantagens propõe-se uma alternativa baseada em superfícies de interacção multi-toque. Pretendemos assim virtualizar as tarefas do DJ, e aliar às vantagens da manipulação de áudio digital uma interacção natural que se apresenta consistente com os gestos tradi-



cionais, identificados através do envolvimento de um grupo de DJs profissionais dispostos a dar o seu contributo para a evolução deste projecto.

## 2. Trabalho Relacionado

Recentemente, foram propostas metáforas alternativas à interacção entre o DJ e o seu *Setup*. Patten et al. desenvolveram o sistema AudioPad [12], uma interface tangível em que um conjunto de objectos físicos permite ao DJ manipular o *playout* dos ficheiros áudio e interligar os diversos componentes. Na mesma linha conceptual surge a proposta de Jordà et. al, a reacTable [9], demonstrando a possibilidade de uma extensa manipulação de trechos sonoros (*loops*) através de objectos tangíveis e multi-toque. Por fim, vale a pena analisar o contributo de Hansen e Alonso na implementação de um conjunto de objectos para *scratch*<sup>1</sup> na reacTable. No entanto, o resultado dos testes efectuados mostrou que o DJ acostumado aos gestos tradicionais teve dificuldade em interagir com os novos objectos [5]. As propostas tangíveis supramencionadas ilustram a possibilidade de novas metáforas perante o paradigma da sequenciação de *loops* áudio ou síntese em tempo real, mas não foram desenhadas para o contexto concreto do DJing, faltando nelas muitos dos requisitos que identificamos no nosso estudo com o grupo de acompanhamento, nomeadamente: controlo por potenciómetros deslizantes e rotativos, acesso directo à mesa de mistura, manipulação directa do áudio e pré-escuta dos canais de áudio. Tais requisitos podem, em última instância, ser traduzidos por mapeamento directo dos conceitos do DJing tradicional nas interfaces virtuais.

Num eixo ortogonal Slayden et al. propõe uma sinergia entre sensores de flexão, óptico e acelerómetro num dispositivo-luva, denominada DjHammer [13]. Também no campo da emulação do *scratch*, Beamish et al. propõe o D'Groove [3], um gira-discos háptico capaz de uma emulação credível, mas que pouco oferece quando comparado a um Híbrido. Em suma, estas novas metáforas apresentam uma maior curva de aprendizagem para o utilizador-DJ, forçando-o a aprender um léxico de gestos não-familiares.

Fora do espectro académico surge um dos sistemas com maior sucesso das últimas décadas - o controlador multi-toque Lemur, Figura 2 - um dispositivo que permite enviar mensagens MIDI/OSC de forma a permitir ao utilizador controlar remotamente o software de DJing. Nas diversas análises que lhe foram dirigidas [4, 1] salienta-se a sua interface modular, composta por *widgets*, que o utilizador escolhe e posiciona no ecrã. Por outro lado, trata-se de um sistema proprietário fechado, não sendo possível ao utilizador/comunidade criar novos objectos. Destinado a uma vasta área de controlo de software de produção musical, não contempla até à data objectos específicos para

<sup>1</sup>Para o leitor não familiarizado, *scratch* denota uma forma de expressão musical realizada num gira-discos [13]. O DJ manipula directamente o áudio, movendo o disco para frente e para trás, complementando com uma acção na mesa de mistura de forma a "abrir" e "fechar" o volume com a outra mão, o que mostra claramente a necessidade de interacção bimanual neste tipo de tarefas.

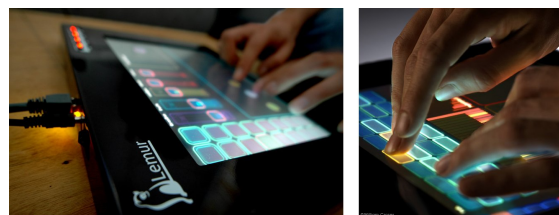


Figura 2. O controlador multi-toque, Lemur. (Extraído de [8])

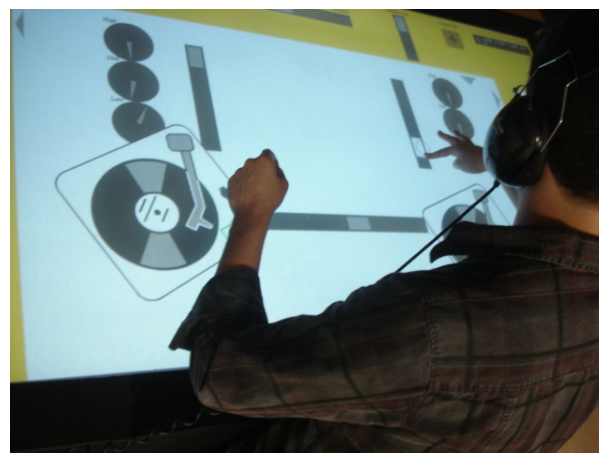


Figura 3. Utilizador testando o MtDjing

as tarefas do DJ (e.g., é notória a falta de um *widget* para simular o comportamento de um gira-discos). Daqui resulta que, apesar de tirar partido da interface multi-toque, o Lemur não oferece aos seus utilizadores DJ uma interacção natural, semelhante aos sistemas tradicionais.

Ainda assim, o Lemur mostra a possibilidade duma utilização proveitosa do multi-toque no contexto do DJing, mas deixa antever alguns desafios de interacção que exploraremos de seguida.

## 3. DJing em Superfícies Multitoque

A actividade de um DJ é muito particular; a análise, realizada durante uma sessão de DJing, permitiu-nos validar um conjunto de requisitos inerentes às suas acções, sendo que as actividades identificadas encontram-se na linha de pensamento de Hansen [6]. A nossa contribuição, MtDjing, posiciona-se nesta linha. Trata-se de uma interface horizontal multi-toque, organizada pelos conceitos fundamentais que encontramos no modelo mental dos DJs: fontes sonoras (gira-discos, sistema digital, leitor de CD), registos áudio (vinil, CD, ficheiro digital), manipuladores de sinal (potenciómetros deslizantes, interruptores, etc...) e as ligações possíveis entre eles. Estes conceitos, claros para o utilizador DJ, são mapeados nos objectos manipuláveis da nossa interface, como ilustrado na Figura 3.

Com o conhecimento adquirido pela análise dos vários *Setups* existentes, compreendemos como o interface para os

DJs pode ter dois papéis vitais: primeiro, reforçando a correlação estreita entre os conceitos virtuais (da aplicação em questão) e os gestos tipificados do DJing; e, em segundo plano, permite-nos desenhar um *workflow* com vantagens face aos *Setups* tradicionais. A escolha do multi-toque, permite-nos beneficiar das seguintes vantagens:

- **Interacção Horizontal**, consistente com os *Setups* analisados, sem recorrer a artefactos/objectos adicionais para interagir com o sistema.
- **Interacção Natural**, a mesma que encontramos nos sistemas tradicionais, despoletada pela acção das mãos/dedos sob objectos do *Setup*, com retorno visual na mesma superfície.
- **Operação Multi-tarefa** caracterizada pela actuação bimanual que oferece vantagens face ao *input* por dispositivos tradicionais, melhorando o desempenho das tarefas do DJ face aos sistemas virtuais [10, 2].

Assim sendo, para virtualizar as actividades dos DJs foi primeiro necessário desenhar interacções compatíveis com os gestos tradicionais, tendo-se recorrido ao grupo de acompanhamento para identificar os gestos típicos para as seguintes acções: play, stop, mistura de dois sinais, interagir com equalizador, fazer pré-escuta do sinal de um canal, alternar entre dois sinais (*fast crossfading*) e alterar volume/velocidade/posição da faixa. Identificou-se que a maioria dos gestos é realizada por acção directa nos componentes hardware (i.e.: potenciómetros, interruptores, disco, agulha), podendo ser mapeada directamente nos componentes "virtuais" da interface. No entanto, existem tarefas mencionadas pelo grupo de acompanhamento que lidam directamente com a preparação/organização do *Setup*, nomeadamente o posicionamento do hardware e interligação dos dispositivos. Estas acções, suscitam alguns conflitos no contexto do multi-toque, pois assumem que é possível efectuar um conjunto de operações no mesmo objecto, usando sempre o mesmo *input* (toque); são elas: mover, rodar, escalar, interligar e "usar o objecto". Para solucionar este problema, optou-se pela aplicação de *handles* nos objectos como proposto por Nacenta [11], permitindo que a interacção dê sempre prioridade às acções específicas do DJing em detrimento do posicionamento dos objectos na cena virtual.

Outro desafio reside no facto da maioria das aplicações de áudio separar o processo de configuração (criar novas pistas, novos objectos DSP, etc...) da interacção com esses objectos, sendo que é uma das desvantagens reconhecidas ao controlador Lemur, que não permite a adição de objectos em tempo real. Para o sistema multi-toque pretendido, o utilizador pode adicionar os seus objectos de DJing e simultaneamente realizar as suas tarefas musicais, como vemos na Figura 3.

Por último, ainda mencionando os desafios da área do DJing, a emulação do comportamento do gira-discos afigura-se como a maior desvantagem identificada para os sistemas baseados em toque, face aos sistemas

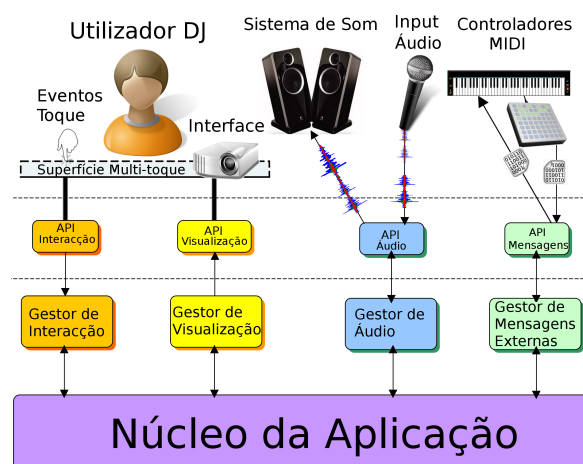


Figura 4. Diagrama da arquitectura proposta.

hápticos [3]. Testar-se-ão diferentes possibilidades com o grupo de acompanhamento de forma a obter a melhor interacção possível com um objecto virtual com a física emulada do gira-discos.

#### 4. Arquitectura do software MtDjing

Uma aplicação tem, neste contexto, de responder a requisitos técnicos ditados tanto pela indústria como pelas expectativas dos utilizadores-alvo; salientam-se assim as medidas de desempenho relativamente à latência percebida pelos utilizadores, capacidade de controlar ou receber mensagens de dispositivos externos<sup>2</sup> e ser independente das escolhas de hardware. As aplicações de áudio profissional operam tipicamente a latências de  $\approx 10\text{ms}$  [3]; na arquitectura MtDjing, ilustrada na Figura 4, o componente de áudio é independente do funcionamento dos restantes, podendo por isso ser executado num ambiente de alto desempenho. A separação dos módulos por funcionalidades e a sua inter-comunicação através do protocolo OSC [14] assegura simultaneamente a interoperabilidade e a possibilidade de um ambiente de execução distribuído, *multithreaded* assente sob *standards* e aberto à mudança.

O sistema encontra-se estruturado em quatro camadas: Núcleo da Aplicação, Gestão, API e Dispositivos. Desta forma os componentes externos interagem com a aplicação através da sua API, e todas as comunicações passam pelo núcleo da aplicação. Este é o módulo principal, responsável por assegurar a comunicação entre os restantes e garantir a persistência dos dados relevantes a cada instante. Enquanto isso, os restantes módulos ocupam-se da gestão de tarefas específicas. O Gestor da Interação interpreta o *input* do utilizador, que mais tarde será encaminhado para o Gestor de Visualização, para que as alterações na interface gráfica sejam consolidadas e o uti-

<sup>2</sup>No DJing, tal como nas restantes áreas da indústria musical, o standard "de facto" é o MIDI [14], embora comecem agora a surgir equipamentos hardware capazes de comunicar no protocolo OSC.

lizador receba *feedback* visual da sua acção. Simultaneamente, as operações no sistema afectam o retorno auditivo, responsabilidade atribuída ao Gestor de Áudio. Este módulo encontra-se implementado em Pure Data, um ambiente de síntese e manipulação de áudio em tempo real usado em projectos como reacTable [9], executando código C para as tarefas de DSP. Por último, a comunicação com dispositivos MIDI/OSC é realizada pelo Gestor de Mensagens Externas, capaz de traduzir as mensagens recebidas, de forma a estender o controlo do sistema, a múltiplos interfaces. Excluindo a camada de visualização, implementada em ActionScript, os restantes módulos foram desenvolvidos em C++, numa abordagem *cross-platform*, restando apenas desenvolver o módulo de comunicação com dispositivos MIDI externos.

## 5. Conclusões e Trabalho Futuro

Apresentámos uma proposta para um sistema baseado em interacção com superfícies multi-toque, capaz de oferecer aos DJs as funcionalidades típicas dos *Setups* tradicionais e as vantagens da virtualização dessas mesmas tarefas. Propomos um modelo conceptual simples, baseado no modelo mental dos utilizadores, obtido pelo envolvimento de um grupo formado por DJs profissionais de diversos géneros musicais. A diversidade dos utilizadores-DJ que acompanharam o desenvolver do sistema trouxe diferentes experiências e necessidades para o projecto, permitindo-nos desenhar um sistema para o qual os conceitos da interface sejam a chave do sucesso. No eixo arquitectural optámos por uma aplicação modular, que facilitará a introdução de novas funcionalidades, optimização do desempenho e manutenção do protótipo.

O trabalho futuro passará pela análise dos testes com novos DJs, de forma a avaliar a proposta e comparar a curva de aprendizagem face às soluções híbridas e virtuais. Por fim, continuar-se-á a adicionar funcionalidade ao protótipo para que este possa ser facilmente equiparado às soluções encontradas no mercado DJ.

## Referências

- [1] Products of interest: Lemur review. *Computer Music Journal*, 29(4):99–112, 2005.
- [2] R. Balakrishnan and K. Hinckley. Symmetric bimanual interaction. In *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 33–40, New York, NY, USA, 2000. ACM.
- [3] T. Beamish, K. Van Den Doel, K. Maclean, and S. Fels. Manipulating music: multimodal interaction for djs. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 327–334, New York, NY, USA, 2004. ACM.
- [4] Philip L. Davidson and Jefferson Y. Han. Synthesis and control on large scale multi-touch sensing displays. In *NIME '06: Proceedings of the 2006 conference on New interfaces for musical expression*, pages 216–219, Paris, France, France, 2006. IRCAM — Centre Pompidou.
- [5] K. Hansen and M. Alonso. More dj techniques on the reactable. In *Proc. of the Conference on New Interfaces for Musical Expression*, 2008.
- [6] K. F. Hansen. *The acoustics and performance of DJ scratching. Analysis and modeling*. PhD thesis, KTH, 2010.
- [7] K. F. Hansen and R. Bresin. *Analysis of a genuine scratch performance*, pages 519–528. LNCS 2915. Springer Verlag, Genova, Italy, selected revised papers, berlin heidelberg edition, apr 2004.
- [8] JazzMutant. <http://www.jazzmutant.com>. consultado em 27/05/2010.
- [9] Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. The reactable: exploring the synergy between live music performance and tabletop tangible interfaces. In *TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 139–146, New York, NY, USA, 2007. ACM.
- [10] Kenrick Kin, Maneesh Agrawala, and Tony DeRose. Determining the benefits of direct-touch, bimanual, and multifinger input on a multitouch workstation. In *GI '09: Proceedings of Graphics Interface 2009*, pages 119–124, Toronto, Ont., Canada, Canada, 2009. Canadian Information Processing Society.
- [11] Miguel Nacenta, Patrick Baudisch, Hrvoje Benko, and Andy Wilson. Separability of spatial manipulations in multi-touch interfaces. In *Graphics Interface*, pages 175–182, Kelowna, B.C., Canada, 2009.
- [12] James Patten, Ben Recht, and Hiroshi Ishii. Interaction techniques for musical performance with tabletop tangible interfaces. In *ACE '06: Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology*, page 27, New York, NY, USA, 2006. ACM.
- [13] April Slayden, Mirjana Spasojevic, Mat Hans, and Mark Smith. The djammer: "air-scratching" and freeing the dj to join the party. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1789–1792, New York, NY, USA, 2005. ACM.
- [14] Matthew Wright, Roger Dannenberg, Stephen Pope, Xavier Rodet, Xavier Serra, and David Wessel. Standards from the computer music community. In *International Computer Music Conference*, pages 711–714, Miami, FL, 2004. International Computer Music Association.



# Combining Different Types of Interaction on Multi touch Surfaces

Tarquínio Mota

CITI, Departamento de Informática, Faculdade de Ciências e Tecnologia,  
Universidade Nova de Lisboa, 2829-516 Caparica, Portugal  
tarquinio@gmail.com

## Abstract

This paper presents a survey of different methods of interaction with single-touch and multi-touch surfaces. While many methods are currently under research, they tend to be difficult to combine. This paper proposes a system that allows the use of different types of interaction on the same surface.

## Keywords

Interface, usability, pervasive, evaluation.

## 1. INTRODUCTION

Touchable interfaces are getting more common every day. Due to the popularization of devices such as PDAs, Tablet PCs, mobile phones and portable music players, touch screens are a technology that most electronic products consumers are getting used to. Initially the applications for these devices were more oriented to use a pen-like device (stylus). However with the appearance of devices such as the iPod Touch, in 2007, touch based interaction became more popular, allowing also the usage of multi-touch gestures. Another feature of the iPod, which is present in many other similar devices, is the presence of an accelerometer. This allows another type of interaction by moving the devices in 3D space.

On small devices, the type of technology used on touchscreens is either capacitive or resistive. When working on a larger scale, these solutions are not viable. The technologies used on smaller displays would be much harder to implement for something as big as a table or a wall, not to mention much more expensive. So there are different approaches that allow to interact in similar way with larger surfaces, using different types of sensors, as will be shown on the next chapter [2-8].

When talking about larger surfaces, the interaction is not only limited to touching the screen. It is also possible to interact by manipulating objects on the surface of the screen. These may be objects with specific visual markers, active objects that can also provide some type of feedback, or simply everyday objects.

In the next section existing prototypes will be presented showing different approaches to multi-touch interaction, focusing specially on surfaces with larger dimensions.

Section 3 presents a new approach to combine different types of interaction and the paper ends with preliminary conclusions and directions for future work.

## 2. RELATED WORK

Most of the approaches to create large scale interactive surfaces involve combining the image from a projector with some way to track the user's hands/fingers. This process is often based on using infrared illumination and capturing a reflection of that light with an infrared camera. Then by processing the captured images it is possible to extract the coordinates.

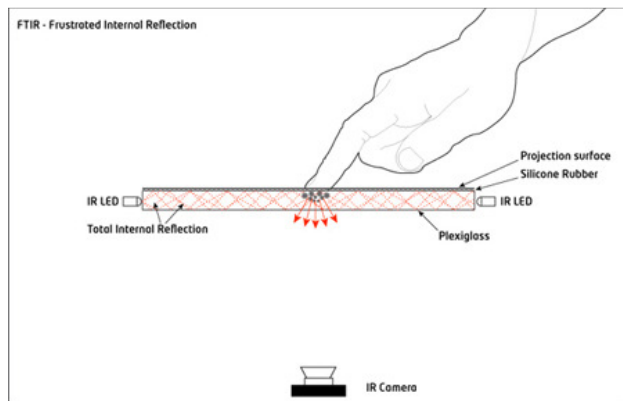
Even though multi-touch technology has been in development for several years [1], the solutions that existed were expensive and the performance was not that good, when compared with the results that are obtained with common devices today.

In 2006 Jeff Han presented a 36 inches wide interactive table with multi-touch sensing capabilities [2]. The main features of the table were high resolution and low cost, while being very scalable in terms of size.

His approach and others that will be presented on this paper are based on the principle of Frustrated Total Internal Reflection (FTIR).

This is a property that certain materials have that allows light to travel within them and always be reflected on the inside. The material commonly used on interactive surfaces is plexiglass.

When an object such as a finger touches the glass the contact area reflects some of the light to the outside. That light is then captured by an infrared camera and that data can be processed to obtain the coordinates.



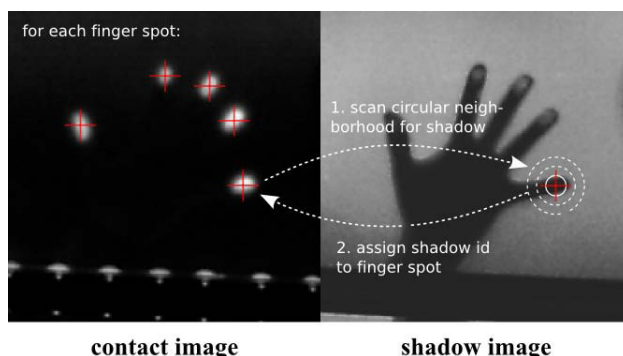
**Figure 1: FTIR touch sensing system**

### 1.1 Interaction using only the hands

Most multi-touch systems are made to detect the fingers only when they are touching the screen. Doing an analogy with a mouse, this means that it is possible to click a specific point by making the pointer jump to a specific spot, it is possible to have a dragging motion (moving the mouse while clicking one button). It is not possible to do the equivalent of just moving the mouse without pressing a button.

Florian et al propose a different system [3] where this is possible. They have two separate sources of light and use them alternately (changing several times per second). This allows the same infrared camera to capture two separate sets of images, one that tracks the objects hovering the screen and the other that tracks the points actually touching the screen (Figure 2).

In this way it is possible to move virtual cursors by hovering the hand over the screen, and detect when each of the fingers are touching the display, simulating a button click.



**Figure 2: Shadow processing [3]**

This system provides some extra information when compared to the previous ones, but some performance is sacrificed in the process. The frame rate of the capture is reduced, since two different sets of images are being

captured, plus changing between them takes time which means some more frames are lost in the process.

Usually when constructing large interactive surfaces the source of the image is a projector, which means the whole system ends up being very large. Wide-angle projectors or a combination of mirrors can be used to minimize the size, but in the end the whole system will still be very large.

Other types of technologies can also be used for touch sensing, like the DiamondTouch [4], that uses electrical signals instead of infrared light in order to identify the users. However this paper will focus mostly on systems using cameras, since they usually require less hardware to develop and are more affordable.

### 1.2 INTERACTION BY MANIPULATING OBJECTS

Another popular way to use interactive surfaces is by manipulating objects on top of them instead of using just the hands.

These objects can be passive and in this case they usually contain a certain type of tag to identify them. They can also be active objects with processing capabilities, sensors or feedback devices. Such objects can measure things like acceleration, distance, position, and offer different feedback types like haptic, visual and audio.

A successful example of this approach is the reacTable [5]. It is an electronic musical instrument, consisting of an interactive table and some objects containing fiduciary markers on them. These markers allow the computer to track and identify objects in the real world, and also identify rotation movements.

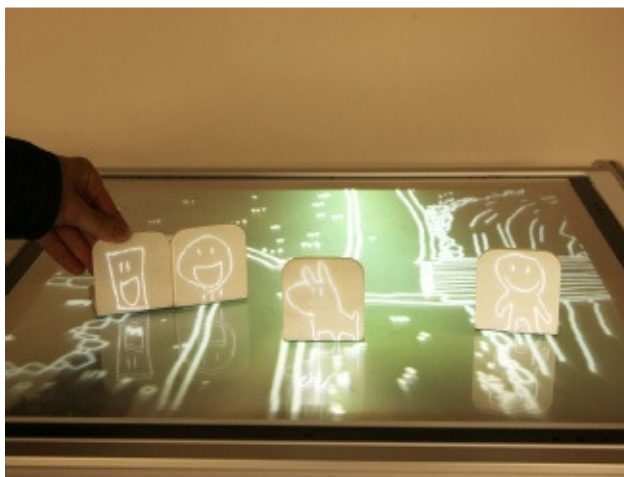
There are different types of objects on the table, and each one has a specific function. By placing objects close to each other links are created between them, changing the output sound wave. Some objects have only one marker, while others have different markers on each face. This means that the same objects can have different functions depending on which face is facing down.



**Figure 3: The reacTable [7]**

Usually these markers need to be touching the table in order to be recognized. Lifting one marker a few

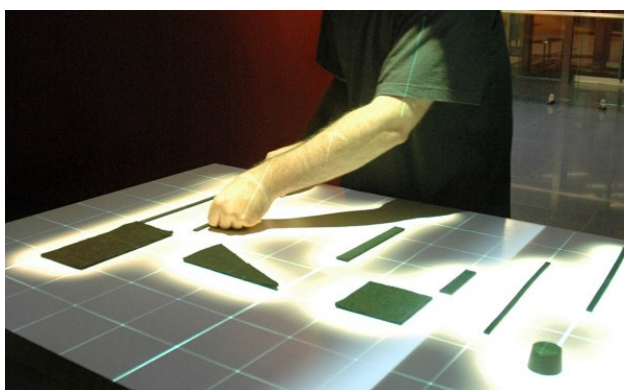
millimeters above the surface will result in a blurred image, unable to be recognized. The reacTable uses the objects only as an input. Other possibility is using the objects both as input and output, as suggested in Tablescape [6]. This work makes the virtual world break out of the screen, by projecting animations on the objects that are placed on the table. In Figure 4 this is used to create an interactive story. The user can move the characters and make them interact by placing them close to each other. According to how the story unfolds, different animations are used for each character.



**Figure 4: Tablescape [6]**

In this case, the objects placed on the screen serve both as input and output.

In 2006 Golan Levin presented Scrapple [7], an interactive table that works as a music sequencer. One interesting thing of this project is that it does not require any special objects or markers; the sequences could be composed by placing random shaped objects on the table.



**Figure 5: Scrapple installation [7]**

By moving, adding or removing objects on the table the users could change the sequence of sounds being played on a loop. The size and shape of the objects also affects how the sound is played.

There is already some work done on combining different types of interaction of the same surface. Leitner et al describe FLUX [8], a multi-touch table/board that can be used with the hands or with a Anoto tracking pen. For the finger tracking they use a normal FTIR approach similar to other described in section 2.1. The only difference is that it is necessary to use a special layer on top on the plexiglass for the Anoto pen to work.

These pens have a small camera aimed at the paper allowing it to identify its position by recognizing small patterns imprinted on the paper. This means that the top layer on this surface must be translucent to have a good projection while being opaque enough for the marks to be visible.

### 3. PROPOSED APPROACH

The goal is to create an interactive surface that combines different types of interaction presented in Section 2. Namely we want to have interaction by using free hands, with a stylus or with objects spread around the table. The detection needs to be as immune to noise as possible since it is meant to be used in public spaces (meaning that there are potentially interferences from other light sources that have to be dealt with).

Combining systems that use FTIR surfaces and fiduciary markers is not possible because the markers have specific patterns that need to be identified, but an FTIR camera will only capture blobs of light. On the other hand, systems that can capture fiduciary markers suffer more interference from normal light, making it harder to track fingers.

The reasons behind choosing a FTIR based system are that it offers a very good price/quality performance, plus it is very scalable and easy to assemble.

Accordingly to these principles, the proposed approach is to use a modified version of an FTIR table that uses active objects with infrared tags. Each object has its own infrared light and is connected wirelessly to the computer. This way the system is totally dependent on infrared light. Each interaction point will be perceived by the system as a blob of infrared light, no matter if it is a finger, a stylus or an object standing on the surface. Instead of using a source of light that is always on, the system will use different sources of light that will be on in different moments in time, blinking several times per second. Since all the lights have a connection to the computer, the software can order the different light sources to blink on request, and then perceive which blobs of light matched the blink order, in order to identify the respective points.

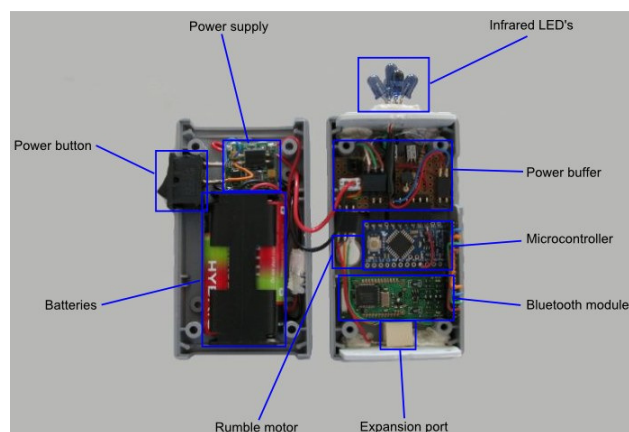
Having active sources of light also helps dealing with the potential interferences to the system. Typically when using infrared systems some type of calibration is used in order to remove interferences. The most common involve capturing the screen with no objects touching it, and then using a background subtraction or similar algorithm to remove the interferences. However this only works well for static

interferences that are already present during the calibration. Changes in room lighting (a light turned on, a flash being fired, or even the movement of shadow/light due to earth rotation) can cause interference to infrared based systems. Some interferences (like the ones caused by the earth moving) can sometimes be dealt with using adaptive background removal. By using active lights it is easier to deal with these situations: the only blobs relevant to the system will blink for sure when requested.

In order to illustrate these mechanisms an example will be presented where a system is designed to track fingers, plus three objects. At a specific moment in time, the software is tracking seven blobs of light. It then orders the lights that illuminate the fingers to blink, and blobs 1, 3 and 5 blink. Then it orders object 1 to blink, and blob 2 blinks. The next step is to order object 2 to blink, and in this case nothing changes. Finally, it orders object 3 to blink, and blob 6 blinks.

After these actions, the system can easily conclude that: blobs 1, 3 and 5 are fingers, blob 2 is object 1, blob 6 is object 3, object 2 is not in range, and blobs 4 and 7 are interferences.

The technology behind the objects to be used will be based on previous work [9]. Some hand controllers were developed that allowed to track and identify the position of different hands in 3D space. The goal here is very similar; we also want to track and identify specific points, so it makes sense to use the same type of hardware, which has proven to work and presents a good performance.



**Figure 6: Hand controller with feedback [9]**

Each object is connected to the computer, enabling the software to request its light to blink on request in order to be identified. Having to wait for the object to blink before it can be recognized causes a slight delay on the detection, but it is a necessary tradeoff in order to get a more trustworthy system.

These objects can be placed on top of the surface or the user can move them hovering over it, and do not require direct contact with the table.

#### 4. CONCLUSIONS AND FUTURE WORK

This paper proposes an alternative to combine different interaction types on the same surface, something that would not be possible with most of the existing approaches for each interaction type.

Having active objects means the system will be more complex but the advantage is that they also offer more possibilities of interaction, like capturing acceleration data or providing different types of feedback.

The system will be implemented and the performance will be compared to other systems with some common features. User tests will be made to evaluate the responsiveness of the users to the combination of these types of interfaces.

#### 5. REFERENCES

1. Lee, S., W. Buxton, and K.C. Smith, A multi-touch three dimensional touch-sensitive tablet, in Proceedings of the SIGCHI conference on Human factors in computing systems. 1985, ACM: San Francisco, California, United States. p. 21-25.
2. Han, J.Y., Multi-touch interaction wall, in ACM SIGGRAPH 2006 Emerging technologies. 2006, ACM: Boston, Massachusetts. p. 25.
3. Echtler, F., M. Huber, and G. Klinker, Shadow tracking on multi-touch tables, in Proceedings of the working conference on Advanced visual interfaces. 2008, ACM: Napoli, Italy. p. 388-391.
4. Dietz, P. and D. Leigh, DiamondTouch: a multi-user touch technology, in Proceedings of the 14th annual ACM symposium on User interface software and technology. 2001, ACM: Orlando, Florida. p. 219-226.
5. Jord, S., et al., The reacTable: a tangible tabletop musical instrument and collaborative workbench, in ACM SIGGRAPH 2006 Sketches. 2006, ACM: Boston, Massachusetts. p. 91.
6. Kakehi, Y., et al., Tablescape animation: a support system for making animations using tabletop physical objects, in SIGGRAPH 2009: Talks. 2009, ACM: New Orleans, Louisiana. p. 1-1.
7. Levin, G., The Table is The Score: An Augmented-Reality Interface for Real-Time, Tangible, Spectrographic Performance. 2006.
8. Leitner, J., et al., Flux: a tilting multi-touch and pen based surface, in Proceedings of the 27th international conference extended abstracts on Human factors in computing systems. 2009, ACM: Boston, MA, USA. p. 3211-3216.
9. Mota, T.M.E., Interfaces based in Gestures and Motion, in Proceedings of Whole Body Interaction Workshop in Conference on Human Factors in Computing Systems 2010, ACM: Atlanta, Georgia, USA.

## **Sessão II**

### **Interacção com Dispositivos Móveis**

Aveiro, 13, 14 e 15 de Outubro de 2010

# Designing Interaction for Outdoor Sports Performance Analysis

João M Brisson Lopes  
Dept. Eng.<sup>a</sup> Informática, IST  
Av. Rovisco Pais 1096-001 Lisboa  
brisson@ist.utl.pt

## Abstract

*This paper addresses the design process of the interactive features required by the users of an application to support athletes and coaches engaged in outdoor sports in analyzing sports performance other than pure fitness. Devices recording position and heartbeat rate allow raw data acquisition for this analysis, but existing software functionality is limited. The paper addresses the collection of interaction requirements for the application to develop that followed a user centred approach with role-playing, interviews, focus groups, scenarios and usability experiments with users involved in outdoor sports practice and competition. The work assessed issues relating to data acquisition, correction, analysis and presentation as well as data persistency and user profiles. A comprehensive set of interaction requirements are then identified and analyzed. Candidate solutions currently being implemented and tested in a full-scale prototype are presented.*

## Keywords

*Interaction, design, sports, performance, data analysis, data acquisition, mobile computing.*

## 1. INTRODUCTION

Developments in outdoor sports data collection devices have increased the range of possibilities of sports performance analysis. This is largely due to price reduction, increased functionality and robustness, and reduced cumbersomeness.

However, the software to analyze and display the huge amounts of data so collected is still a long way from satisfying user needs beyond simple physical fitness analysis.

Cardiac data acquisition and recording has gone from mere instantaneous heartbeat rate display to sampled at regular times heartbeat rate recording and transmission. Simultaneously, wrist-held GPS devices, which record time, position and altitude, have merged into one single device with heartbeat rate recorders. This enabled outdoor sports practisers to gather more precise and comprehensive data, thus opening the way for better data analysis of sports practice and competition.

Equipment vendors know that data upload from devices and basic data display are not enough to differentiate themselves in the market and thus bundle some software for data analysis and storage with devices sold. However, this software is limited because it is usually given away with the devices (or available from the Internet) and thus cannot support the wide (and, sometimes, incompatible) full range of requirements set by different outdoor sports.

Some vendors provide sites where users can register, store data and make some (reduced) visualization and analysis. Some sites allow data sharing between users,

and thus set virtual communities. But each site operates only with data produced by devices from a given vendor. Thus, end users have to do with limited capabilities software that may even not accommodate their basic requirements.

Furthermore, end users in informal sports groups or at club level, who do not use devices from the same vendor, are unable to share data and make comparisons other than the most basic.

The above shows the need to reassess the issues involved and propose a new design for an interactive application supporting sports performance analysis.

### 1.1 Objectives

The objectives of the work reported here were:

- Analyze significant devices available on the market and their software.
- Identify interaction and user requirements that were not met by existing software.
- Discuss issues related to interaction requirements identification and solution design.

### 1.2 Contents

First, this paper addresses the methodology that was adopted, and then analyzes data recording devices available on the market and the software available with them. This is followed by a discussion on data, raw track data and user data. Data correction is then addressed. Finally, the main features' requirements are presented with results and conclusions.



## 2. METHODOLOGY

A user centred methodology was followed because most users, although knowledgeable in the specifics of the sports they practice, have little to no knowledge of sports data collecting devices and how to process the data collected. The application to develop had to be simple to use and its added value had to be made clear to users.

### 2.1 First Phase

The approach started with a series of projects of the User Centred Design course at IST carried out by students as the main course assignment. Starting with identical objectives, students were given the freedom to conduct the project as they saw fit and explore the solution domain at will. For this they conducted inquiries, interviews and experiments with users, individually and from focus groups. Cultural probes ([Gaver99], [Anderson04]) were also used. This involved some 30 users (amateur athletes, athletes, coaches and MDs) in various phases.

Students explored the problem domain from devices to collect data to the operation of the application, searching for alternative designs, following the methodology of [Sharp07], and collected usage scenarios.

The result was a set of requirements tentatively implemented in several prototypes that were explored with users. The tests allowed prototypes to evolve and be tested with users in controlled usability experiments.

The main results from this phase were:

- A set of requirements for the functionality of the application to develop
- A set of interactive solutions to support user requirements

Without knowing it, the students involved in the projects were part of the user experiment themselves because they were allowed to try any approaches and solutions without a definite background. Observation of student activities reporting in class seminars was very important because it allowed us to pinpoint difficulties and pitfalls in the process of the application development and testing later on.

### 2.2 Second Phase

The second, and current, phase had the objective of producing a working prototype and testing its components. This involved data collection on the field, data upload, conversion, data storage, data filtering and correction and analysis.

With the help of the members of the National Kids Ski Team we collected data at the 2010 La Scara International Championships and at the National Alpine Ski Championships (Slalom) and the Spanish National Veterans Alpine Ski Championships. Others contributed with data from cross-country and jogging activities.

The tests involved getting raw data in various formats into the application together with maps (both from graphics files or from interactive sources, e.g., Google Maps) and produce visualizations to be interactively used in data validation and analysis.

Of course, this involved testing with users, with special focus on the steps of getting data to the application, data correction and data display and manipulation.

### 2.3 Next Phases

The current phase will be followed by two more phases: platform selection and final implementation and testing.

The current prototype has been implemented for MS Windows in C#. This choice of platform prevents the application to be used in other popular operating systems like Mac OS, a limitation that is common to the software of most device vendors.

Furthermore, portability means supporting the solution for graphical data display on a portable graphical system or package, like OpenGL [Woo99] or GTK [GTK10], with user interface support, and the selection of a programming language, like C++. These options must be evaluated on the face of the exact functionality requirements set by the results of the second phase prototype experiments.

The final phase will involve building and testing the final version with users.

## 3. DEVICES AND ASSOCIATED SOFTWARE

Data for sports performance analysis originates from devices carried by athletes while practicing or competing.

Vendors usually provide software for data upload to a Windows PC that, most of the times, does some data processing and shows textual summaries and data charts.

Next, we briefly present a few significant devices from some device vendors (Suunto, Polar, Garmin and Geonate) and their software.

### 3.1 Suunto

Suunto was one of the first suppliers of wrist-held GPS devices. Suunto X9i [Suunto06] and, recently, X10i, record time, position, altitude (both GPS and altimeter) and speed at user preset intervals, but no heartbeat rate. This is a basic instrument whose minimum 10 seconds recording rate makes it more appropriate for slow speeds or long distances.

Both devices connect to Windows based PCs through a serial USB connection. They come with a Google Earth track upload program that gets the list of tracks stored in the device and, on user selection, creates a temporary KML<sup>1</sup> file, and fakes the Web download the file to Google Earth. The KML temporary file holds only track points' locations.

All data is only available with Suunto Trek Manager, a program that accesses the device and uploads and stores track data. This program can export track data in the KML, SDF<sup>2</sup> and TXT (Waypoint Plus) formats.

Trek Manager requires calibrated maps in common image formats (BMP, JPEG or GIF) to show tracks on Map View. Profile View shows altitude and/or speed versus time or distance. Track View shows data sum-

---

<sup>1</sup> Keyhole Markup Language.

<sup>2</sup> Suunto Data Format.



maries as well a table with coordinates of all track points. These can be eliminated but not edited.

There are no user profiles and all data and maps are stored in a treelike directory structure.

### 3.2 Polar

Polar RS800 [Polar] with Polar G3 GPS pod and Polar ProTrainer software (Windows only) make up a very complete set aimed at physical fitness. Unlike most devices from other vendors, the RS800 connects to a Windows PC through an infrared IrDA USB adapter.

The software has complete user support, with very comprehensive user profiles and training logs and diaries, and fitness analysis, but this is totally built around heartbeat rate recordings and focused on fitness.

The program's charts show heartbeat rate, altitude, speed, cadence and power against either time or distance. Data error correction is restricted to the definition of the ranges of admissible values. Track sections definition or deletion is done as in audio or video track editing and force users to know start and end times or distances because no map view supports more advanced editing.

There is no Google Maps support, but the program can export data to Google Earth or save data to GPX<sup>3</sup> files.

### 3.3 Garmin

Garmin supports combined GPS and heartbeat devices with a lightweight application, Garmin Training Center, that runs on Windows and Mac OS and supports data upload through a serial USB connection. Data can then be uploaded to Garmin Connect, a website providing more advanced data analysis, user support and data sharing.

Both Garmin Training Center and Connect export data to Google Earth, GPX files and TCX<sup>4</sup> files, a GPX extension supporting timed-stamped heartbeat rate data [Garmin08].

Charts on both products display heartbeat rate, speed, altitude and altitude change as functions of distance, time since recording start or (unique to Garmin) absolute time.

However, data correction can only be made at the level of lap deletion (a lap is a very broad concept that may mean the track between two locations set by pressing a device button).

Garmin uses its vector GPS IMG maps to display tracks with Training Center and Google Maps with Connect. This allows Training Center to operate offline, while Connect can use Google Maps because it is already a Web application.

Data collection is done either at one-second intervals or every time the user changes direction (Smart Recording). The earlier has the disadvantage of filling up device memory too fast (little more than 3 hours), while the later may miss recording relevant data if the user travels on a straight line for a long time.

<sup>3</sup> GPS Exchange format.

<sup>4</sup> Training Center Database XML.

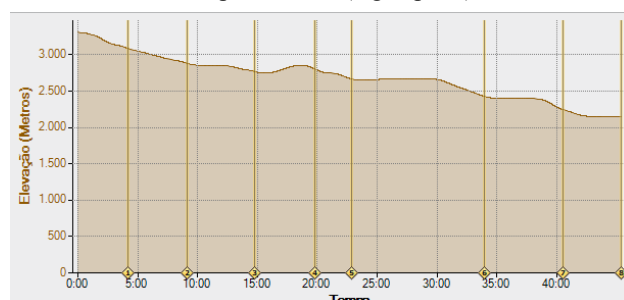
### 3.4 Geonaute

Geonaute KeyMaze [KeyMaze] is a line of GPS wrist-held devices. The top of the line device (Keymaze 700) also records heartbeat rate at a user selected sampling rate that can be as low as one second. Data is uploaded through a serial USB connection to two Windows applications, KeyMaze 500-700 Utility and Geonaute Software.

Both applications rely on Google Maps to display tracks logged with devices. The Utility has no concept of user profile or activity type and displays only simple charts of altitude, speed and heartbeat rate versus time. Conversely, Geonaute Software has extensive user profiles and activity types and descriptions and its charts plot altitude, speed, heartbeat rate, pace and power against time or distance while allowing for other chart enhancements (see Figure 1). Both applications provide textual data (summaries), with Geonaute Software providing the most extensive and comprehensive.

However, there is no support for data correction and the comparison of data from different runs is rather limited on the Software and nonexistent on the Utility.

The Utility exports data to KML (and Google Earth), CSV<sup>5</sup>, enhanced with heartbeat rate values GPX, NMEA RMC<sup>6</sup> and TXT files. GPX and TXT files contain all data; the other formats miss some of the parameters (e.g., speed). The Software exports only to KML, GPX and Fitlog formats and, as it is more physical fitness oriented, does not include all parameters (e.g., speed).



**Figure 1: Altitude change with time. Speed is not shown; equal distance vertical lines provide indirect speed information.**

### 3.5 Remarks on Devices and Software

The above shows different solutions for the analysis and display of data collected by outdoor sports oriented devices and the software bundled with them.

#### 3.5.1 Chart Representations

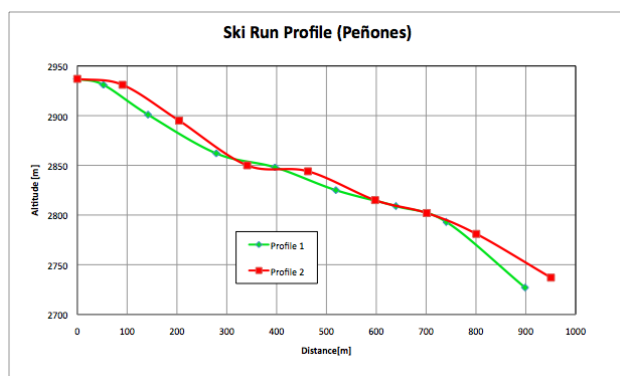
In general, solutions target physical fitness objectives and completely forget the analysis of performance especially in sports played against the clock (e.g., mountain bike downhill, skiing) or where acrobatic aspects are relevant (e.g., wakeboarding).

<sup>5</sup> Comma Separated Values, easily read by spreadsheet applications.

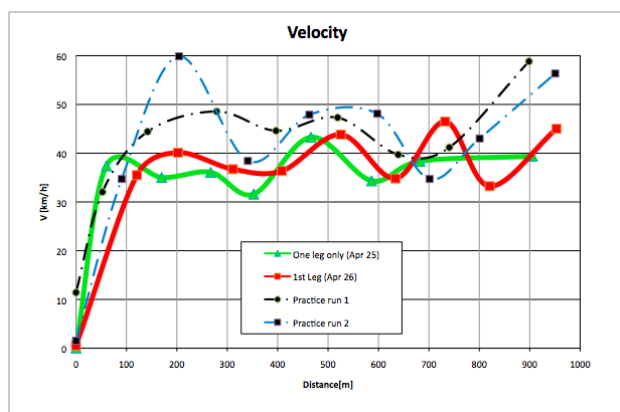
<sup>6</sup> National Maritime Electronics Association, Recommended Minimum data for GPS: position, velocity and time.

Chart representations invariably come down to the plot of one or more parameters versus either time or distance. There is no provision for more advanced correlations, like speed and/or heartbeat rate versus hill grade.

Figure 2 (altitude-distance) and Figure 3 (speed-distance) charts illustrate the above. All software displays any of the two charts, but only one at a time. There is no way to visually correlate speed with hill grade and how speed changes with hill grade. This would not be easy to plot on a speed and height versus distance chart since we would end up with a cluttered chart.



**Figure 2: Altitude-distance approximate profiles of a giant slalom ski run.**



**Figure 3: Velocity-distance profiles of training and competition runs on the ski run of Figure 2.**

### 3.5.2 Section Identification

The concept of track sections and section-by-section analysis is rather vague. Most software define laps, but the concept can vary widely from laps around a race track, to passing repeatedly close to the same location (although very different courses may be followed in between passages) to a given time interval (every  $x$  seconds) or distance (every  $y$  meters). Furthermore, users must either signal lap completion by pressing a button on the wrist-held device (impossible to do while ski racing or wakeboarding) or accept whatever the software thinks a lap or a section is or has been set on the device.

The giant slalom ski run shown on Figure 2 has three different sections that must be differently skied. Therefore, it would be convenient to be able to analyze each section separately in order to better understand the run and improve timings. This shows the need for a feature

enabling users to interactively identify sections on a map, chart or data table.

Moreover, as in the case of Figure 1, we may need to assess the same track in two ways: the whole track or a special section of the track that begins 29.5 minutes from the start and ends some 5 and a half minutes later. This section was easily identifiable on a Google Maps view but it demanded some 15 minutes of data manipulation with a text editor and a spreadsheet application to get the approximate start and end times. This was because the application did not allow pointing to a location on the map view and immediately show it on the height-time chart or on the track points' table.

### 3.5.3 Data Pruning and Correction

The identification of sections would also enable discarding irrelevant data from data sets. This is the case of data collected while waiting for the start of a race, after the race (while the device is not stopped) or in between races (device not paused).

Data correction also needs the point-to-map-to-select-data to point to data that may contain errors (position, altitude, heartbeat rate) and start correction procedures (see section 5.2).

### 3.5.4 Data Comparison

To compare data from two or more passes that followed the same track one needs to be able to make apart the parameters that make the differences between the passes, especially in the cases of almost identical passes.

This is because there may instances when two racers are separated by fractions of a second after a couple of minutes or kilometres. Existing software would create map views and charts where data from one racer would be visually indistinguishable from that of the other.

Therefore, the software to develop must be able to detect such cases and revert to representations where track lines are artificially separated or, if coincidental, there is additional graphical information that clearly shows the little differences between passes or racers.

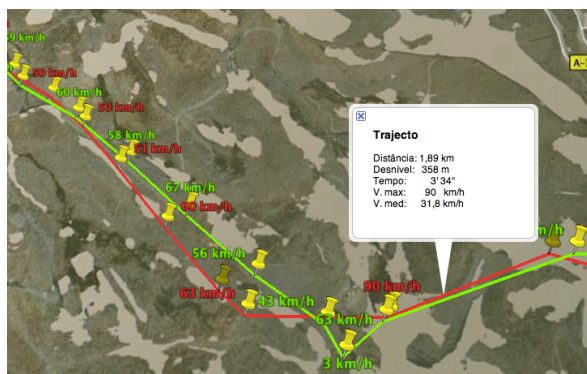
### 3.5.5 Data Enrichment

Sometimes one needs to publish or make a permanent representation of some track data (e.g., for informational or educational purposes). This is the case of Garmin software that produces Google Earth path animations.

However, it is much more important to be able to associate data to track visualization and turn it on and off. Figure 4 shows a Google Earth display of a KML file that was enhanced with speed value labels that can be turned on and off. Callouts display section summaries.

### 3.5.6 Multimedia Content

Suunto Trek Manager is the only software mentioned here that allows associating a position on a track with images and notes. All other applications lack this feature that can be of advantage, especially when images and video of different runners have been shot and may show errors or reasons for different performances.



**Figure 4: Data enriched tracks, displaying speed at selected track points by extensive manual enhancement of a Suunto Trek Manager generated KML file.**

#### 4. DATA

When asked what data to collect, users mentioned that the relevant parameters were: heartbeat rate, position, speed, calories burned and hill grade. Some mentioned altitude as next most relevant, followed by, in decreasing order of relevance, climb/descent rate, ambient temperature, blood pressure and sweating rate. Except once, no users ever mentioned data sampling rates as an issue.

In general, users confused parameters measured by devices with parameters displayed by them but computed from the earlier (e.g., hill grade was often mentioned but altitude was not) and forgot time most of the time.

Almost all users mentioned that they would like to be able to choose the data units to be used. They emphasized that choice of data units for speed, hill grade and climb rate were essential.

However, users never mentioned choice of position coordinates units as an issue, thus showing that the majority is not aware of the existence of many such units (UTM, geographic units or others) or that many maps use local geographic coordinate systems and grids instead of a unique global coordinate system. This is because users are much more familiar with applications like Google Maps or Bing Maps than they are with survey maps.

##### 4.1 Data Parameters Requirements

Given the above, the application to develop had to handle data of the following parameters (the actual parameters depending on the device used):

- Time (and date)
- Position
- Speed
- Altitude
- Heartbeat rate
- Blood pressure
- Sweating rate

Also, users should be able to select the units for each kind of parameter.

In what regards position coordinate systems, the application should use a global system with widespread use. The choice came naturally to the WGS84 datum. Any data in other systems should be transparently converted to this datum, especially maps (e.g., Portuguese military grid

maps) and users must be warned of that when loading new maps into the application.

#### 4.2 Data Ownership and Sharing

Users made clear that data from each run should be associated with a specific user (its owner) and that it could only be shared with other users designated by the user, including granting access to coaches.

Sharing data was mentioned as important because it would bring in the social aspects of sports. Also, in many individual sports practiced in teams, there is the need to share information between team members to improve performance (first runners usually relay course information to team members running after). Most of the time there is the need to compare team members' runs.

Therefore, users needed to be able to grant/deny access on single and groups of runs data to single users or groups of users at any time. There was no mention of standing automatic access grants.

Users stressed that access granting must not be transferable.

#### 4.3 User Profiles

To associate data with users (i.e., assign ownership) and keep and check access to other users' data, there must be a minimal user profile. There were opposite user opinions as to whether users must login or not, but the issue was settled when access granting was mentioned.

Furthermore, some personal data is needed to interpret run data. This is the case of indicative heartbeat rate values, especially maximum heartbeat rate and target heartbeat rate ranges for different types of sports practice.

Also relevant is the need to classify new data according to type of sport. A short list of the sports practiced by a given user is always better to handle than a long list of all possible sports. This classification brings in the benefit of a more logical structure for the set of run data of a given user and also helps in analyzing the data because parameters are not the same for all outdoor sports.

#### 5. DATA VALIDATION AND CORRECTION

Most interviewed users and users from focus groups assumed that data collected by devices would always be correct. However, this cannot be assumed, especially in the case of position and altitude.

##### 5.1 Sources of Data Errors

Position determination by GPS devices greatly improved after removal of the Selective Availability feature on May 1<sup>st</sup>, 2000, but errors remain. GPS devices need to receive signals from at least 4 satellites to unambiguously determine position. Natural obstacles (caves, mountains and trees), man-made obstacles (buildings and power lines) and even persons block signal reception. Proximity to water may also produce spurious readings.

Receiving antennae size is also a factor: hand-held GPS devices usually produce more accurate positioning than wrist-held devices. The latter may even produce incorrect speed values on sudden arm movements (e.g., in a fall to the ground)



Figure 5 shows typical GPS position errors in the vicinity a power line in a narrow valley.



**Figure 5: GPS position errors due to an overhead power line and mountain cluttered sky (all locations should have fallen inside the white area).**

Altitude reading errors depend on device type. GPS altitude errors are at least double than position errors or higher. Altimeters require calibration before use or end up producing errors on the order of 50 m or more.

Altitude errors may not be as dramatic as position errors since, most of the time, analysis does not require correct altitude values but altitude differences where errors cancel out. However, some cases requiring correct altitude values still remain (e.g., start and end altitudes of a ski run for a race results list).

Heartbeat rate is not immune to errors. Heartbeat rate is usually acquired by fitting a sensor strap to the athlete's chest. Heartbeat rate values are continuously transmitted to the receiving device by small range weak radio signals. Tests done on a treadmill with two different recording devices show that devices usually display the same heartbeat rate but one is faster than the other to detect changes.

Errors occur when the sensor strap slips from its correct fitting position because there is too much sweat between the skin and the strap or the strap elastic fit loosens.

In both cases, heartbeat readings show a slowdown of heartbeat rate (because the strap is away from its correct position close to the heart). If there is no filtering, analysis may end up concluding that the athlete's performance has greatly improved.

Such events may be easily detected because there is a sudden heartbeat rate decrease that may go all the way down to rates close to or lower than the athlete's heartbeat rate at rest. Another sign is that this happens while there is no change to other parameters (e.g., speed remains constant). Also, the change is so sudden that it cannot be confused with the athlete stopping the exercise and going into recovery where a slowdown of heartbeat rate extends over several minutes in a decaying pattern.

## 5.2 Data Errors Correction

The above shows that the application interface must support tools to correct / eliminate the following main raw data error types:

- Position errors
- Altitude errors
- Heartbeat Rate errors

### 5.2.1 Position errors correction

Previously, we have seen that most software bundled with devices allows only two position error correction options: simple deletion of the location point or track section (or lap) where the error occurs.

When it is possible to change location coordinates, speed should also be changed because location and speed are interrelated. Moreover, software from device vendors has seldom the capability of directly point to and select a location on a map to edit its parameters<sup>7</sup>. None has the capability to detect locations where position or speed might be wrong.

This generated several requirements. First, the user should be able to select a location on a course, be it on a course map, a table of track points or on a chart of course parameters. If more than one of these views are simultaneously present on the application screen, selection of a location on any one of them must immediately be shadowed by highlighting of the same location on all others.

Second requirement was that the user could displace any location by a drag-and-drop operation on a course map. Naturally, all other representations of the dragged point should be updated accordingly.

The third requirement concerns automatic detection of location errors on the assumption that sharp turns are unnatural unless there is a pause in the course. Sharp turns may be detected by assuming that course tracks are smooth and any point not close enough to an assumed smooth track is a candidate point for inspection.

Lastly, once a candidate point has been selected for position correction, a hint (based on the smooth course track assumption) for the new parameters of the location should be displayed.

### 5.2.2 Altitude errors correction

Correction of altitude errors can be made against analogical or digital references, on the assumption that location errors are negligible or have been corrected.

While altitude correction against digital references poses no major problem if interpolation is handled with care, altitude corrections against analogical references (i.e., maps showing altitude contour lines) is a different matter.

One way is to correct the altitude of a small number of locations, followed by the interpolation of altitude correction to the remaining locations.

For a single location, the procedure involves selecting a set of points that surround the location and are over altitude contour lines on a contour map, and letting the software carry out altitude interpolation using those points.

For exact altitude values, Web resources like Google Elevation API may be used.

<sup>7</sup> The same requirement applies when selecting points for data values enrichment to be (selectively) displayed with course plots (e.g., on Google Earth, see Figure 4).

### 5.2.3 Heartbeat rate errors correction

Heartbeat errors correction is the most difficult to perform since there is no reference to get correct values from. The simplest way is to do away with the wrong data and assume that heartbeat changes smoothly with time and that sudden changes are errors. Provided that exercise effort remains unchanged, data can be assumed as back to normal when heartbeat rate returns to levels close to those registered prior to a sudden drop.

## 6. MAIN REQUIREMENTS

Given the above, the most relevant requirements for an application for outdoor sports practice performance analysis and evaluation can now be presented.

- Application to run without Internet connection. The application must run without connection to the Internet and display track plots to allow it to be used on the field. Track plots will be superimposed on a map if appropriate maps exist in the system and the user selects to use them.
- Three data view modes. Track data will be displayable in three modes: map view, chart view and table view. All three views may not be simultaneously displayed on the screen. More than one chart may be displayed at the same time.
- Map view, chart view and data view will display all data selected by the user irrespectively of whether it is the data of a single raw data set or section or a group of raw data sets or sections.
- Data comparison. When comparing almost identical data, map view, chart view or data view shall use display enhancements to make out or visibly amplify differences between the data sets that are being compared. On map view, a track may be placed parallel to its original location or visualization enhancers used.
- Progress views. When comparing data to assess progress, the application shall compare data sets in chronological order and highlight differences.
- Zoom and pan on maps and charts. Zoom and pan will be available in map and chart views.
- Choice of parameters to display on charts. Users will be able to select which parameters to display on a chart from all parameters in the parameter set.
- Location selection. When more than one view is presented on the screen, selecting a location in one of them will highlight the same location on the other views.
- Selection on map view. The location of a track point on a map closest to the pointer will be always the candidate location for selection. Callouts showing relevant track point data parameter values may be used at user's choice.
- Weather information (user supplied) will be attached to each new raw data set. Sections will inherit such weather information.
- Section identification. The user will be able to identify one or more (contiguous or not) sections in a track. Sections may be assigned visible or invisible status or be deleted. The user

will be able to classify complete tracks or sections as good or bad examples and support this classification with multimedia content (see below).

- Error correction. The user will be able to edit the properties of locations or remove locations from a track data set. Editing will follow the guidelines set in section 5.2.
- User defined locations. The user will be able to add, modify and delete user-defined locations (markers) and associate them to a track data set to document events such as mistakes, accidents, etc. The user shall be able to associate multimedia content to marked locations (see next).
- Multimedia support. Users shall have the option to associate multimedia content (text, notes, images and video) to any locations and user-defined markers. Sections will inherit multimedia content from parent sections / raw data.
- Extendable parameter set. In order to allow application expansion to support user defined new parameter types (both raw or derived), the parameter set must be extendable by user customization and a tool should be available to help define new parameters and their units.
- Data structure, persistency and protection. The application shall keep all data in a treelike structure and regularly back it up. The user must doubly confirm data deletion and the application shall keep a copy of the deleted data for a given time set by user preferences.
- Data import. The applications will only import data from files produced by the programs released with the devices, thus committing the task of data upload from devices to vendors' applications.
- Data export. The application shall be able to export a full data set, a section, a group of data sets or a group of sections selected by the user in the extended file formats (GPX or TCX) and formats suitable for publication (CSV, tabbed text files, HTML tables).

## 7. CONCLUDING REMARKS

The methodology followed to determine interaction requirements needed by an application for the analysis of outdoor sports performance proved to be adequate and up to the task at hand. In fact, it was possible to identify many essential interaction requirements that are overlooked by software bundled by vendors with devices for data recording.

This was the outcome from adopting several complementary techniques of addressing users and user needs: interviews, role-playing, user groups, user focus groups, scenarios, usability testing and surveys.

The extensive set of interaction design requirements reported earlier has already produced results that are shown on Figure 6 to Figure 9. These results depict several solutions for the interface that were tested with users.

Work is progressing on the full-scale prototype. It is already possible to load data and maps and evaluate solutions for the interaction requirements (see Figure 9).



Figure 6: A solution for track display on a map and a chart, featuring a callout with data of a selected track point.

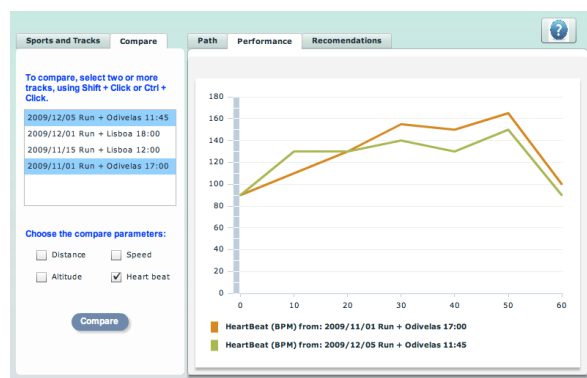


Figure 7: Solution for data selection and simple comparison of data from two runs on the same track course.

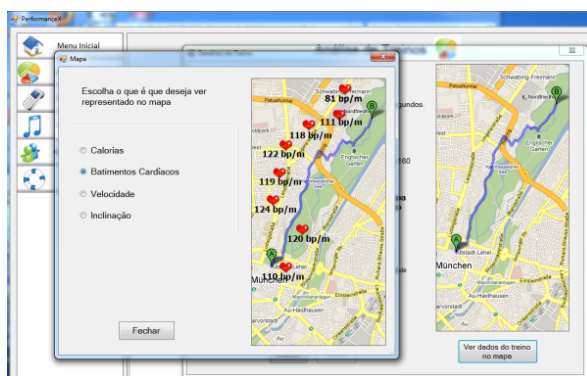


Figure 8: Solution for data enhancement on map view, showing heartbeat rate at selected locations.

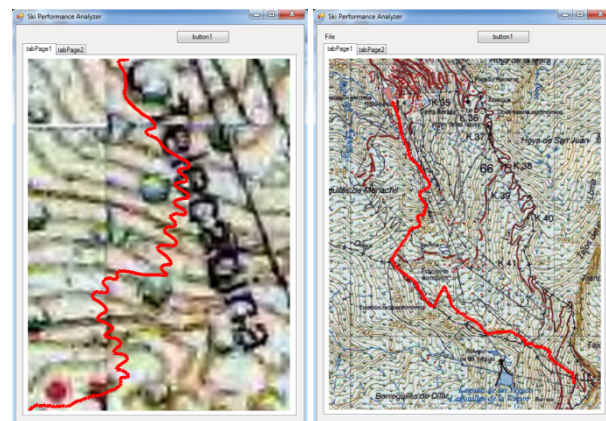


Figure 9: Data selection tests where zoom and pan on the track of Figure 1 on a map (right) allow the identification and isolation (left) of an important section.

## 8. ACKNOWLEDGEMENTS

The author wishes to thank the (Portuguese) Winter Sports Federation (FDI) and the members of the National Kids and Veteran Ski teams.

Thanks also the IST MSc students involved in the early studies: Carlos Fonseca, Francisco Carapuça, Frederico Freitas, Hugo Damas, João Benedito, José Lopes, Kristjan Jurisalu, Pedro Gonçalves, Rafael Kuffner, Sarah Vilarinho and Tomaz Kuklis.

Special thanks go to Nuno Gonçalves for data and the many discussions.

## 9. REFERENCES

[Anderson04] Anderson, A, Dittlau, A, Funkel, K, Probing Students, *CADE2004 Web Proceedings of Computers in Art and Design Education Conference*, 29 June – 1 July 2004.

[Garmin08] Forerunner 205/305 Owner's Manual, <[http://www8.garmin.com/manuals/984\\_OwnersManual.pdf](http://www8.garmin.com/manuals/984_OwnersManual.pdf)>

[Gaver99] Gaver, W., Dunne, A. Pacenti, E.. Cultural Probes, *Interactions: New Visions of Human-Computer Interaction*, vol. 6, No 1, 21-29. 1999.

[GTK10] The GTK+ Project, <<http://www.gtk.org>>.

[KeyMaze] Geonaute Keymaze 300-500-700, <<http://keymaze.geonaute.com>>

[Polar] Polar ProTrainer 5, <[http://www.polar.fi/en/products/training\\_software/polar\\_protrainer\\_5](http://www.polar.fi/en/products/training_software/polar_protrainer_5)>

[Sharp07] Sharp, H, Rogers, Y, Preece, J, Interaction Design (Beyond Human-Computer Interaction), 2nd ed., John Wiley & Sons, 2007.

[Suunto06] Suunto X9i User's Manual, Suunto Oy, 2006.

[Woo99] Woo, M, Neider, J, Davis, T, Shreiner, D, OpenGL Programming Guide, 3<sup>rd</sup> ed., Addison-Wesley, 1999.

# DetiGuide: Interagindo com um ecrã de grandes dimensões através de um dispositivo móvel

Rui Palha

Departamento de Electrónica, Telecomunicações e Informática (DETI), Universidade de Aveiro  
ruipalha@ua.pt

Paulo Dias  
IEETA – Universidade de Aveiro  
paulo.dias@ua.pt

Rui A. Costa  
Ubiwhere Lda  
rcosta@ubiwhere.com

Beatriz Sousa Santos  
IEETA – Universidade de Aveiro  
bss@ua.pt

## Sumário

A interacção com sistemas multimédia interiores ou exteriores baseados em ecrãs de grandes dimensões ainda tem limitações e problemas a ser superados. Os mais recentes dispositivos móveis, equipados com sensores (tais como acelerómetros, bússolas e ecrãs sensíveis ao toque), podem ser uma alternativa para interagir com estes ecrãs já que esta solução não implica *hardware* adicional (o utilizador pode usar o seu próprio telemóvel).

Neste trabalho apresentamos as diferentes opções e os passos que levaram à criação do DetiGuide: um guia informático para o Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro que permite aos utilizadores interagir com a informação apresentada num ecrã instalado no átrio do departamento através de um dispositivo móvel *Android*. O principal objectivo deste protótipo é estudar o uso de dispositivos móveis para interagir com ecrãs de grandes dimensões.

## Palavras-chave

Computação Móvel, Ecrãs, Interacção Humano - Computador, *Android*

## 1. INTRODUÇÃO

Os mais recentes dispositivos móveis integram novas tecnologias que potenciam possibilidades de interacção com ecrãs de grandes dimensões. O objectivo deste trabalho é estudar e adaptar estas tecnologias (acelerómetros, bússolas digitais e ecrãs sensíveis ao toque) disponíveis nos mais recentes dispositivos móveis para encontrar novas formas de interacção. Foi desenvolvido o DetiGuide: um guia informático para o Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro que permite duas formas distintas de interacção aos utilizadores de modo a obter informações apresentadas num ecrã de grandes dimensões através de um dispositivo móvel *Android*. Ao longo deste artigo são introduzidos alguns projectos de investigação relacionados com o trabalho realizado, a apresentação de um estudo comparativo da usabilidade das alternativas de interacção desenvolvidas, assim como as principais conclusões do trabalho realizado.

## 2. TRABALHOS RELACIONADOS

Existem já alguns protótipos que utilizam dispositivos móveis para interagir com ecrãs de grandes dimensões. Alguns dos protótipos proporcionam diversas formas de interacção. Henze *et. al.* [Henze08] realizaram uma comparação experimental entre quatro técnicas de interacção

(*touching, pointing, scanning* e *user-mediated object interaction*), sendo recolhidos dados com utilizadores em diferentes cenários. Foram implementados quatro protótipos: um sistema de interacção móvel em ambientes inteligentes, um guia turístico móvel, um guia de museu móvel e um protótipo para a interacção com cartazes de publicidade móvel. Rukzio *et. al.* [Rukzio07] realizaram um estudo sobre a interacção dos dispositivos móveis com o mundo real. A utilização da câmara dos dispositivos móveis foi explorada por Ballagas *et. al.* [Ballagas05], que introduziram duas técnicas de interacção: *Point & Shoot* permite aos utilizadores seleccionar objectos utilizando códigos visuais para criar um sistema de coordenadas absoluto sobre a superfície da tela; *Sweep* permite a utilização de um dispositivo móvel como um rato óptico com vários graus de liberdade. Jiang *et. al.* [Jiang06] utilizam a câmara do dispositivo móvel para controlar o cursor de um ecrã. Erbad *et. al.* [Erbad08] desenvolveram o *MAGIC Broker*, uma ferramenta de interacção com o público, usado em ecrãs de grandes dimensões, permitindo a interacção com ecrãs por voz, enviando SMS ou utilizando um browser através do dispositivo móvel. Hardy *et. al.* [Hardy08, Hardy09] desenvolveram *Touch & Interact*, uma técnica de interacção em que um dispositivo móvel toca um ecrã em qualquer posição de modo a realizar selecções. Rahman *et. al.* [Rahman09] realizaram um estudo para tentar averiguar



as potencialidades do uso do acelerómetro para interacção nos dispositivos móveis.

A comunicação entre dispositivos móveis e ecrãs de grandes dimensões pode ser efectuada utilizando vários métodos de comunicação (infravermelhos, marcadores visuais, *Bluetooth* [Cheverst05], *GPRS*, etc.). O *C-Blink* [Miyaoku04], que é um marcador de luz e a técnica *Flashlight* [Shirazi09], que permite uma interacção detectando com uma câmara o flash do telemóvel, utilizam fontes luminosas para efectuar a comunicação.

A colaboração entre utilizadores é também um aspecto importante na interacção entre dispositivos móveis e ecrãs de grandes dimensões. Loregian *et. al.* [Loregian06] conceberam e implementaram um protótipo que permite a vários utilizadores em simultâneo interagir com um ecrã de grandes dimensões através de dispositivos móveis com comunicação infravermelhos. Zhong *et. al.* [Zhong09] desenvolveram o *Doodle Space*, uma aplicação interactiva desenvolvida para ecrãs de grandes dimensões que permite a vários utilizadores pintarem em conjunto utilizando a câmara de dispositivos móveis. Vajk *et. al.* [Vajk08] desenvolveram o *TiltRacer*, um jogo desenvolvido para ecrãs de grandes dimensões, que é controlado pelos sensores de movimentos de um telemóvel, e que permite a vários utilizadores jogar em simultâneo, através da comunicação *Bluetooth*. Leikas *et. al.* [Leikas06] conduziram um estudo sobre a interacção social na utilização de serviços multi-utilizadores em dispositivos móveis.

Através dos exemplos referidos, foi possível verificar que uma interacção entre dispositivos móveis e ecrãs de grandes dimensões é caracterizada pelas formas de interacção disponíveis, o tipo de comunicação implementado e o nível de interacção proporcionado aos utilizadores.

### 3. DETIGUIDE

O DetiGuide é uma aplicação baseada no *site* do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro. Foi desenvolvida para permitir o uso de um telemóvel *Android* para interagir em espaços públicos com um ecrã de grandes dimensões e obter informações variadas (localizações, contactos, actividades). O principal objectivo deste protótipo é testar diferentes formas de interacção possíveis com uma configuração incluindo um telemóvel *Android* e um ecrã de grandes dimensões controlado por um computador, como ilustrado na Figura 1.



Figura 1 - Elementos da configuração usada na DetiGuide e comunicação entre estes e o utilizador.

O DetiGuide proporciona dois modos de interacção: navegação através das páginas seleccionando uma das quatro setas esquerda/direita/cima/baixo (ver Figura 2) e

selecção (por exemplo seleccionar um nome na página da lista de contactos). Um menu e dois botões virtuais são disponibilizados no ecrã do telemóvel para comutar entre os dois modos e para as duas operações principais: Selecção e *Home*. O DetiGuide pode ser configurado para usar como dispositivo de entrada tanto o acelerómetro/bússola digital ou o ecrã sensível ao toque de acordo com os seguintes movimentos:

- Acelerómetro/Bússola Digital: a navegação é feita rodando o dispositivo móvel para cima, baixo, esquerda ou direita para se deslocar para a página correspondente. A selecção é feita usando a metáfora de uma bola sobre um plano que se move de acordo com a gravidade e a inclinação do telemóvel: pequenos movimentos do dispositivo móvel resultam em movimentos do cursor no ecrã;



Figura 2 - Navegação entre páginas movendo o telemóvel: rodar para a esquerda, direita, cima e baixo irá seleccionar a página correspondente.

- Ecrã sensível ao toque: para navegar entre as páginas, o utilizador move o dedo na direcção correspondente. A selecção é feita usando o ecrã para mover o cursor de uma forma similar ao *touchpad* de um computador portátil, mas apenas utilizando a posição inicial do dedo no ecrã e a última posição antes de levantar o dedo do ecrã, devido a limitações do dispositivo móvel.



Figura 3 - Selecção de itens utilizando o ecrã como um *touchpad*.



As figuras 2 e 3 mostram um utilizador no modo de navegação usando o acelerómetro e a bússola digital e no modo de selecção usando o ecrã sensível ao toque, respectivamente. Os botões virtuais mencionados acima podem ser vistos no ecrã do dispositivo móvel.

A figura 4 mostra alguns dos menus que compunham a aplicação *Android* que controla o DetiGuide.

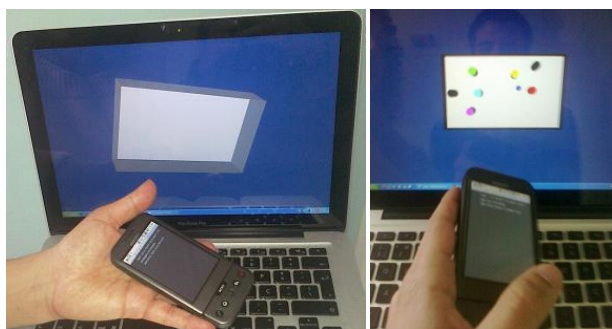


**Figura 4 - Imagens da aplicação *Android* que controla o DetiGuide: menu principal (esquerda) e menu de escolhas de interações (direita).**

#### 4. FERRAMENTAS UTILIZADAS

O sistema operativo móvel utilizado é o *Android*, visto pela comunidade científica como umas das mais promissoras tecnologias emergentes na área da computação móvel [Speckmann08]. A flexibilidade é a sua maior vantagem, já que é uma plataforma móvel *open-source* totalmente personalizável. Neste trabalho, foi usada a versão 1.6 do *Android*, a versão mais recente disponível no início do trabalho.

O desenvolvimento foi feito em diferentes fases. Em primeiro lugar, uma simples demonstração foi criada para sincronizar o movimento do dispositivo móvel com um paralelepípedo mostrado no ecrã. Num segunda fase, um jogo mais complexo foi desenvolvido, que consistia em controlar uma bola através de movimentos feito com o dispositivo móvel (Figura 5).



**Figura 5 - Sincronização entre o dispositivo móvel e um paralelepípedo (esquerda); jogo (direita).**

Esta demonstração usa os dados do acelerómetro para controlar a bola e os dados da bússola digital para torná-lo independente da orientação do dispositivo. Estes protótipos, usados para compreender as capacidades dos sensores e os protocolos de comunicação, conduziram ao protótipo final do DetiGuide.

O dispositivo móvel utilizado foi o *HTC Dream*, o primeiro dispositivo móvel a incluir o sistema operativo *Android*. É equipado com um acelerómetro de três eixos (*Bosch Sensortec BMA150*, com uma taxa de actualização de 3 kHz) e uma bússola digital de três eixos (*Asahi Kasei AK8973*). Relativamente às comunicações, o telemóvel disponibiliza 3G, Wi-Fi e *Bluetooth*. Pelo facto do *Bluetooth* não estar totalmente implementado no sistema operativo utilizado, a comunicação foi implementada através da arquitectura cliente-servidor. Inicialmente, foi testado o protocolo TCP/IP; no entanto, após testes iniciais, verificou-se que este protocolo não era o mais adequado, já que as suas características causavam atrasos significativos. Este problema foi resolvido com a utilização do protocolo UDP, resultando numa comunicação mais rápida e melhorada.

Para o desenvolvimento de *software* foram utilizados dois emuladores para facilitar a programação sem usar o dispositivo real: *Android SDK*, um emulador de um dispositivo móvel *Android* e *Sensor Simulator* para simular os dados fornecidos pelo acelerómetro e bússola digital. Como linguagem de programação foi usado Java na plataforma móvel, a única linguagem disponível para *Android*. Java e C# foram seleccionados para desenvolver os protótipos que correram no computador ligado ao ecrã. Bibliotecas adicionais, como *VTk (Visualization Tool Kit)*, foram também usadas em alguns exemplos. As plataformas de desenvolvimento foram o *Eclipse* e o *Visual Studio 2008*.

#### 5. TESTES COM UTILIZADORES

De modo a validar as opções feitas e avaliar as interações disponíveis no DetiGuide, foram realizados testes com utilizadores. Foi elaborada uma lista de tarefas para observar a forma como o utilizador reagia a cada uma delas. As tarefas foram escolhidas com base na importância que desempenham no contexto do DetiGuide. As quatro tarefas utilizadas, perguntas cujas respostas podiam ser encontradas pesquisando o DetiGuide, eram as seguintes:

1. Em que ano foi fundado o Departamento?
2. Qual o telefone de contacto de um determinado docente?
3. Em que ano foi atribuído ao Instituto um dado estatuto?
4. Qual o código no Portal Académico de uma dada disciplina?

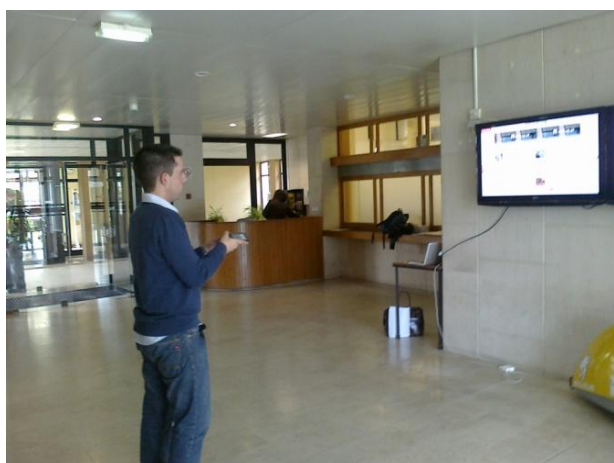
Duas das tarefas implicavam que o utilizador usasse apenas as interações com o acelerómetro/bússola digital, enquanto nas outras duas tarefas só poderia utilizar as interações com o ecrã sensível ao toque. Metade dos utilizadores utilizaram as duas formas de interação

(*touch* e acelerómetro) numa dada ordem e a outra metade em ordem inversa para se poderem compensar e estudar eventuais possíveis efeitos de aprendizagem.

Enquanto na primeira e na terceira tarefa apenas era necessário a utilização do modo de navegação, a segunda e a quarta tarefa exigiam o uso do modo de navegação e de selecção. Na segunda tarefa era ainda necessária a utilização do *scroll* dado a lista de docentes não ser toda visível numa só página.

Os utilizadores não tiveram a possibilidade de experimentarem o sistema antes de efectuarem os testes, tendo sido efectuada uma demonstração das interacções disponíveis no início de cada teste. Esta decisão prendeu-se com o facto de tentar averiguar se existiu uma aprendizagem na utilização dos modos de interacção. No final dos testes, foi pedido aos utilizadores que respondessem a um questionário recolhendo a sua opinião quanto a alguns aspectos do DetiGuide. Também foram recolhidos dados pela própria aplicação, tais como o tempo total de cada tarefa e o número de erros por tarefa.

Os participantes nos testes foram sobretudo alunos, mas também professores. Foram realizados testes com um total de 51 utilizadores em 2 sessões, 26 na primeira sessão e 25 na segunda sessão, compreendidos numa faixa etária entre os 19 e 46 anos, com uma percentagem de 96% dos utilizadores do sexo masculino e 4% do sexo feminino. Esta amostra, embora não seja representativa do público em geral, é muito representativa dos possíveis utilizadores do DetiGuide. Os testes realizaram-se no átrio do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro no dia 20 de Abril e 24 de Maio de 2010. Foi utilizado um dos ecrãs disponíveis no local, um computador pessoal e um telemóvel *HTC Dream*. Na Figura 6 é possível visualizar um utilizador navegando no DetiGuide.



**Figura 6 - Utilizador navegando no DetiGuide.**

Para o registo do grau de dificuldade das tarefas observado pelo observador e indicado pelos utilizadores foram usadas escalas de *Lickert* de 5 níveis, com valores entre 1 e 5 (1 Muito Difícil e 5 Muito Fácil). Relativamente às formas de interacção, os utilizadores manifestaram a concordância ou discordância em relação às afirmações apresentadas, onde 1 representa Discordância Total e 5 Con-

cordância Total. O Grau de Satisfação dos Utilizadores na utilização de cada uma das formas de interacção foi indicado também com valores entre 1 e 5 (1 Nada Satisfeito e 5 Muito Satisfeito).

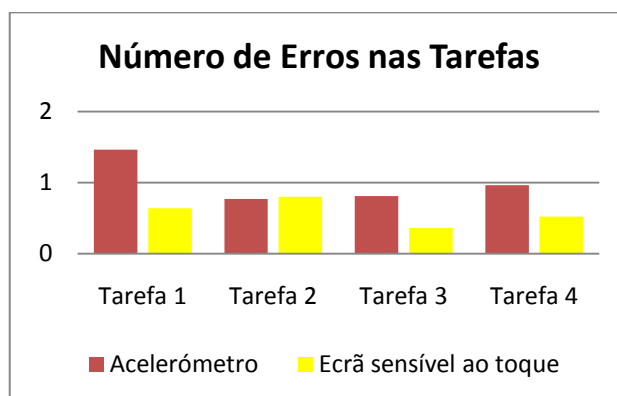
## 6. RESULTADOS OBTIDOS

Esta secção apresenta as principais conclusões dos testes com utilizadores realizados ao DetiGuide. Alguns utilizadores na primeira sessão de testes queixaram-se de que a selecção de itens no ecrã era pouco confortável devido ao tamanho reduzido do botão de selecção. Devido a este facto, o tamanho do botão de selecção foi aumentado da primeira para a segunda sessão de testes. Essa alteração teve influência nos resultados da segunda sessão de testes, comprovada pela diminuição do tempo das tarefas, em especial aquelas em que era necessário utilizar o botão de selecção (Tarefas 2 e 4), e também pela opinião dos utilizadores que atribuíram à Tarefa 4, uma das afectadas pelo tamanho do botão de selecção, um grau de dificuldade menor.

O tempo total e o grau de dificuldade das tarefas que foram afectadas na segunda sessão de testes por esta alteração, serão analisados separadamente dos da primeira sessão de testes, enquanto os que não foram afectados serão analisados em conjunto.

### 6.1 Número de Erros das Tarefas

O número de erros nas tarefas foi contabilizado partindo do princípio que um erro consiste em cada ecrã desnecessário visualizado para realização da tarefa. Com base na Figura 7, podemos verificar que os utilizadores realizaram um maior número de erros utilizando o acelerómetro/bússola digital do que com o ecrã sensível ao toque, com excepção da Tarefa 2, o que reforça a ideia de que os utilizadores tiveram maiores dificuldades utilizando o acelerómetro/bússola digital. Podemos constatar também que foi efectuado um número menor de erros na realização da tarefa 2 usando o acelerómetro/bússola digital do que na tarefa 1, apesar da tarefa 2 possuir um grau mais elevado de dificuldade, o que sugere que houve uma rápida aprendizagem por parte dos utilizadores na utilização da interacção através do acelerómetro/bússola digital. Este facto poderá ser justificado por esta tarefa ser bastante intuitiva.



**Figura 7 - Número de erros nas tarefas dos testes com utilizadores.**

## 6.2 Análise do Tempo das Tarefas

O tempo total das tarefas refere-se ao tempo desde que o utilizador saiu da página inicial do DetiGuide até ao instante em que chegou à página que continha a informação necessária para responder à pergunta. Com base nas Figuras 8 e 9, podemos verificar que as tarefas 2 e 4 levaram mais tempo em média em relação às tarefas 1 e 3, tal como era esperado, pois implicavam navegação e selecção. A tarefa 2 foi a mais demorada pois além de ser necessária a utilização do modo de navegação e selecção como na tarefa 4, era também necessário o uso do *scroll*. Os utilizadores demoraram mais tempo a executar as tarefas com o acelerómetro/bússola digital do que com o ecrã sensível ao toque, com excepção da tarefa 4, o que sugere que os utilizadores tiveram maiores dificuldades na utilização do acelerómetro/bússola digital. Da primeira para a segunda sessão de testes, verifica-se uma ligeira diminuição do tempo médio total das tarefas, sobretudo na tarefa 2.

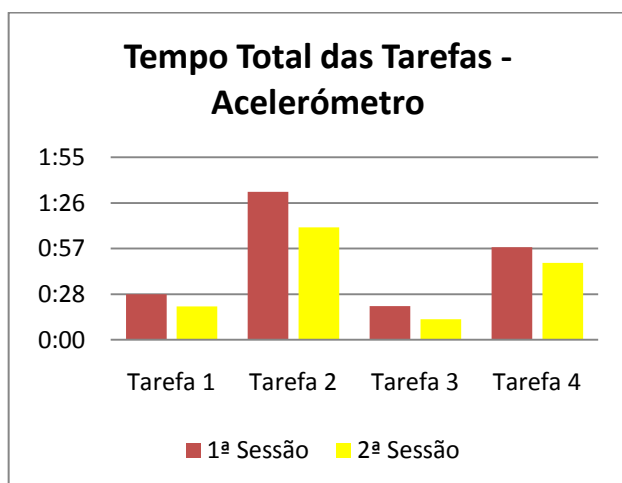


Figura 8 - Tempo médio das tarefas nos testes com utilizadores utilizando a interacção através do acelerómetro/bússola digital.

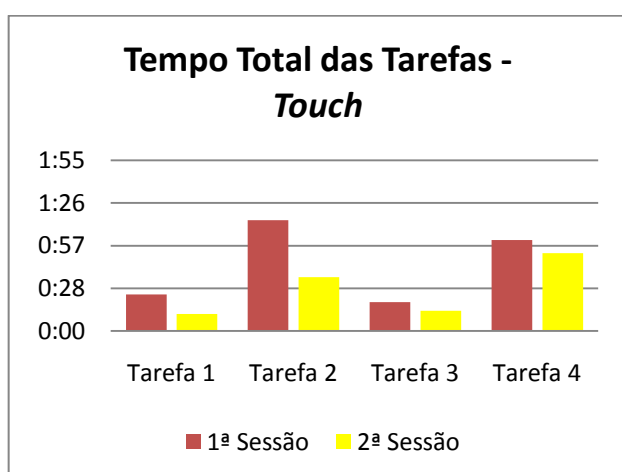


Figura 9 - Tempo médio das tarefas nos testes com utilizadores utilizando a interacção através do ecrã sensível ao toque.

## 6.3 Grau de Dificuldade das Tarefas

Como referido, cada utilizador indicava no final de cada tarefa o grau de dificuldade. Com base nas Figuras 10 e 11, podemos verificar que as tarefas 2 e 4 foram consideradas mais difíceis do que as tarefas 1 e 3, como era esperado. Os utilizadores acharam a primeira tarefa mais fácil através do ecrã sensível ao toque e na tarefa 3 o nível de dificuldade experimentado pelos utilizadores foi o mesmo em ambas as interacções. Relativamente à tarefa 2 os utilizadores acharam que na primeira sessão de testes o nível de dificuldade foi maior através da interacção com o acelerómetro/bússola digital, e na segunda sessão de testes o nível de dificuldade experimentado pelos utilizadores foi o mesmo. Na tarefa 4 os utilizadores acharam que foi mais fácil a interacção através do acelerómetro/bússola digital.

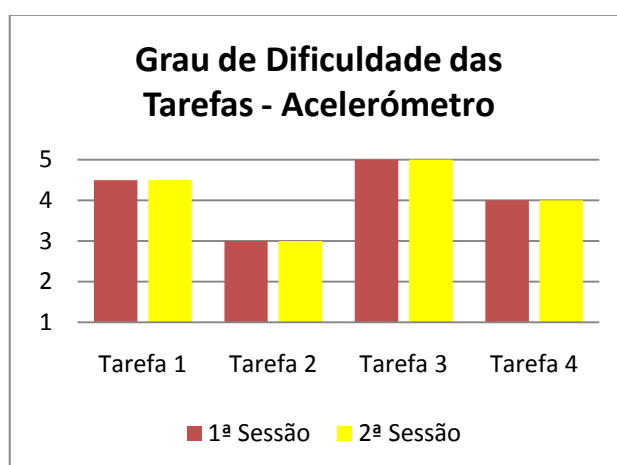


Figura 10 - Grau de dificuldade das tarefas nos testes com utilizadores utilizando a interacção através do acelerómetro/bússola digital.

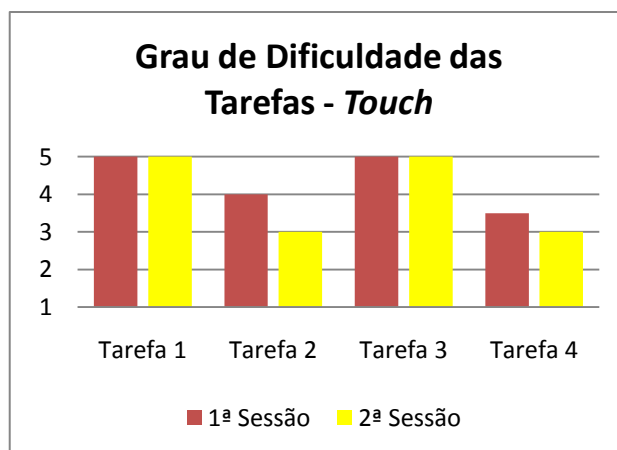
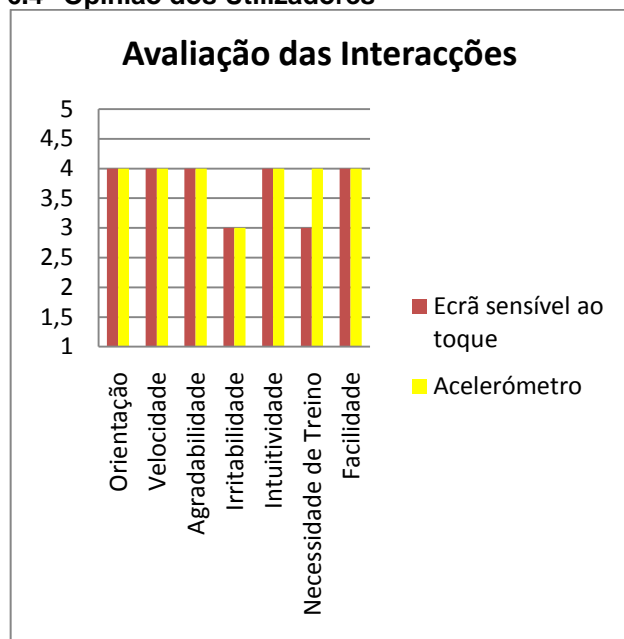


Figura 11 - Grau de dificuldade das tarefas nos testes com utilizadores utilizando a interacção através do ecrã sensível ao toque.

## 6.4 Opinião dos Utilizadores



**Figura 7 - Resultados dos questionários dos utilizadores após a realização dos testes.**

A Figura 12 sugere que os utilizadores acharam que as duas formas de interação disponíveis eram bastante usáveis e equilibradas entre si. É de realçar o facto de os utilizadores acreditarem que a interação através do acelerómetro/bússola digital necessitar de mais treino do que com o ecrã sensível ao toque, o que parece natural visto que a utilização do acelerómetro/bússola digital é ainda pouco explorada, enquanto o ecrã sensível ao toque é uma tecnologia bastante divulgada no mercado de dispositivos móveis.

## 6.5 Satisfação e Preferência dos Utilizadores

Os utilizadores consideraram as duas formas de interação disponíveis equilibradas em termos de satisfação, assim como ambas apresentaram uma boa usabilidade. Apesar de equilibrados, a interação através do ecrã sensível ao toque foi a preferida dos utilizadores, com 27 utilizadores a preferirem esta interação contra 24 que preferiram o acelerómetro/bússola digital. Esta conclusão justifica-se já que os utilizadores cometeram um maior número de erros e demoraram mais tempo para realizar as tarefas com o acelerómetro/bússola digital.

## 6.6 Discussão

As aplicações desenvolvidas ao longo deste projecto foram o controlo de um paralelepípedo, um jogo e o *DetiGuide*. Todas estas aplicações são controladas através dos movimentos efectuados com o dispositivo móvel *Android*. O controlo do paralelepípedo permitiu aprender a usar e a aceder aos diversos sensores do dispositivo móvel utilizado, e também testar a arquitectura de comunicação implementada, especificamente quanto à sua velocidade e a fiabilidade. Numa fase posterior, a criação do jogo teve como objectivo testar o potencial do acele-

rómetro para interação. Por último, o propósito do *DetiGuide* foi testar diferentes formas de interação possíveis usando um dispositivo móvel.

Numa segunda fase, o desenvolvimento da plataforma *DetiGuide* permitiu desenvolver um protótipo facilmente adaptável para uma utilização em locais públicos.

Dos testes com utilizadores ao *DetiGuide*, foram retiradas algumas conclusões, fundamentadas tanto pela análise dos dados recolhidos como pela opinião dos utilizadores. As principais conclusões foram as seguintes:

- A interação através do ecrã sensível ao toque causou alguma irritabilidade junto dos utilizadores pelo facto de este sistema não ter um comportamento similar ao de um *touchpad* de um computador portátil. Esta interação apenas utilizava a posição inicial do dedo no ecrã e a última posição antes de levantar o dedo do ecrã. Devido a limitações do dispositivo, não foi possível alterar esta característica do sistema;
- O movimento efectuado com o telemóvel durante a interação através do acelerómetro/bússola digital possuía o eixo vertical invertido, o que foi considerado pouco intuitivo por parte de alguns utilizadores, já que é necessário deslocar o telemóvel para baixo para seleccionar uma opção na parte de cima do ecrã e vice-versa. Contudo, analisando os resultados dos testes, a maioria, 63%, concordou com a atribuição efectuada aos movimentos;
- Na primeira sessão de testes, quando foi utilizada a opção de selecção com o acelerómetro/bússola digital, o controlo do cursor foi considerado pouco confortável por parte de alguns utilizadores, devido ao tamanho reduzido do botão de selecção. Na segunda sessão de testes, o botão de selecção foi aumentado, passando a ocupar cerca de metade do ecrã do telemóvel. Apesar do tempo total das tarefas ter diminuído nas tarefas afectadas pela utilização do botão de selecção, o espaço para manipulação da interação através do ecrã sensível ao toque diminuiu consideravelmente. Apesar de ser esperado que esta alteração afectasse a opinião dos utilizadores em relação à irritabilidade e a agradabilidade, tal não se verificou;
- Na interação através do ecrã sensível ao toque, os utilizadores acharam que o cursor deveria permitir maior precisão do movimento. No entanto, devido ao espaço disponível no ecrã para efectuar esta interação, principalmente depois de aumentar o botão de selecção, o aumento da precisão iria aumentar consideravelmente o número de movimentos necessários que o utilizador teria de efectuar até conseguir atingir o item pretendido, o que levaria a uma maior irritabilidade por parte do utilizador, assim como o aumento do tempo total das tarefas.

Na medida em que o trabalho desenvolvido foi exploratório, algumas opções foram tomadas com base no tempo disponível para a sua implementação, como é o caso da plataforma escolhida para o DetiGuide, sendo à partida conhecida a sua pouca portabilidade.

## 7. CONCLUSÕES E TRABALHO FUTURO

Apresentou-se o DetiGuide, um protótipo desenvolvido para testar as potencialidades de utilização de um telemóvel *Android* para interacção com um ecrã de grandes dimensões num local público. Foram usadas duas formas de interacção que exploram o ecrã sensível ao tacto e o acelerómetro e bússola digital do telemóvel, tendo-se realizado testes com utilizadores.

Como referido, a sincronização do dispositivo móvel com o sólido e o desenvolvimento do jogo foram os primeiros passos dados no sentido de utilizar as capacidades dos sensores do telemóvel *Android*, assim como a comunicação entre o telemóvel e o ecrã. A principal conclusão a reter é que a escolha do protocolo usado para estabelecer a comunicação entre o telemóvel *Android* e o servidor do computador tem de ser cuidadosa devido ao facto de poder comprometer o desempenho do sistema.

Quanto à plataforma *Android*, apesar de necessitar de bastante tempo de aprendizagem, demonstrou o seu potencial durante a realização do trabalho desenvolvido, mostrando ser uma opção para dispositivo de entrada de um sistema multimédia interactivo.

Ao contrário do que se pensava inicialmente, a interacção através do acelerómetro/bússola digital superou bastante as expectativas, mostrando ser uma opção para interacção a utilizar em futuras aplicações. No entanto, a escolha das técnicas de interacção a usar tem que ser muito bem pensada, de acordo com as características das aplicações nas quais vão ser integradas. Terá que existir uma relação de compromisso entre os vários parâmetros a considerar, como o tamanho dos botões da interface e a adequação da sensibilidade dos sensores.

Existem alguns aspectos que poderão ser melhorados para aumentar a qualidade do software desenvolvido:

- Incluir na opção de selecção da interacção através do acelerómetro/bússola digital um movimento “inteligente” na vizinhança de *links*, que consiste em colocar o cursor a descrever uma trajectória automática na direcção do objecto mais próximo do movimento efectuado com o telemóvel;
- Realizar uma outra experiência com utilizadores, testando a precisão do acelerómetro e do ecrã sensível ao toque, de modo a tentar encontrar soluções para melhorar o controlo de cursor do DetiGuide através dos sensores do telemóvel. Para isso, poder-se-ia adaptar o jogo desenvolvido durante este projecto, de modo a que o movimento da bola também pudesse ser controlado com o ecrã sensível ao toque, e avaliar o desempenho do jogo comparando as duas formas de interacção;
- Tornar o DetiGuide uma plataforma portátil, de modo a facilitar a introdução da informação na aplicação, assim como torná-la acessível de qualquer local. Espera-se que novas versões sejam desenvolvidas utilizando plataformas Web, aumentando significativamente a portabilidade do sistema, podendo esta ser acedida de qualquer local com ligação à Internet e a partir de qualquer sistema operativo;
- De modo a aumentar a versatilidade da aplicação desenvolvida para *Android*, para utilizar a aplicação num local público, será necessário adaptar o software desenvolvido para as várias plataformas móveis que existem no mercado, podendo o sistema operativo do telemóvel ser detectado através de *Bluetooth*, aquando da transferência da aplicação para o telemóvel. Para esta situação, a utilização da versão mais recente do sistema operativo *Android*, versão 2.1, poderia permitir ao servidor da plataforma informar o portador do dispositivo móvel que pode descarregar uma aplicação para o seu telemóvel de modo a interagir com o ecrã.

## AGRADECIMENTOS

Os autores agradecem à *Ubiwhere* Lda a cedência do dispositivo móvel e a disponibilidade demonstrada no acompanhamento do trabalho.

## REFERÊNCIAS

- [Ballagas05] Ballagas R., Rohs M., Sheridan J. (2005), 'Sweep and point and shoot: phonecam-based interactions for large public displays', *CHI'05 Extended Abstracts Human Factors in Computing Systems* (Portland, USA: ACM), pp. 1200-1203.
- [Cheverest05] Cheverest K., Dix A., Fitton D., Kray C., Rouncefield M., Sas C., Lagoudakis G., Sheridan J. (2005), 'Exploring bluetooth based mobile phone Interaction with the hermes photo display', *Proceedings of the 7th International Conference on Human Computer Interaction with mobile devices and services* (Salzburg, Austria: ACM), pp. 47-54.
- [Erbad08] Erbad A., Blackstock M., Friday A., Lea R., Al-Muhtadi J. (2008), 'MAGIC Broker: A Middleware Toolkit for Interactive Public Displays', *Proceedings of the Sixth Annual IEEE International Conference on Pervasive Computing and Communications* (IEEE Computer Society), pp. 509-514.
- [Hardy08] Hardy R., Rukzio E. (2008), 'Touch & Interact: Touch-based Interaction of Mobile Phones with Displays', *Proceedings of the 10th international conference on Human-Computer Interaction with mobile devices and services* (Amsterdam, The Netherlands: ACM), pp. 245-254.



- [Hardy09] Hardy R., Rukzio E., Wagner M., Paolucci M. (2009), 'Exploring Expressive NFC-Based Mobile Phone Interaction with Large Dynamic Displays', *Proceedings of the 2009 First international Workshop on Near Field Communication*, pp. 36-41.
- [Henze08] Henze N., Broll G., Rukzio E., Rohs M., Zimmermann A. (2008), 'Mobile Interaction with the real world', *Proceedings of the 10th international conference on Human-Computer Interaction with mobile devices and services* (Amsterdam, The Netherlands: ACM), pp. 563-565.
- [Jiang06] Jiang H., Ofek E., Moraveji N., Shi Y. (2006), 'Direct Pointer: Direct Manipulation for Large-Display Interaction using Handheld Cameras', *Proceedings of the SIGCHI conferece on Human Factors in computing systems* (Montréal, Canada: ACM), pp. 1107-1110.
- [Leikas06] Leikas J., Stromberg H., Ikonen V., Suomela R., Heinila J. (2006), 'Multi-User Mobile Applications and a Public Display: Novel Ways for Social Interaction', *Proceedings of the 4th Annual IEEE International Conference on Pervasive Computing and Communications* (IEEE Computer Society), pp. 66-70.
- [Loregian06] Loregian M., De Michelis G., Martini P. (2006), 'Directional Interaction with Large Displays Using Mobile Phones', *Proceedings of the 4th annual IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 196-200.
- [Miyaoku04] Miyaoku K., Higashino S., Tonomura Y. (2004), 'C-blink: a hue-difference-based light signal marker for large screen interaction via any mobile terminal', *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology* (Santa Fe, USA: ACM), pp. 147-156.
- [Rahman09] Rahman M., Gustafson S., Irani P., Subramanian S. (2009), 'Tilt Techniques: Investigating the Dexterity of Wrist-based Input', *Proceedings of the 27th International Conference on Human Factors in Computing Systems* (Boston, USA: ACM), pp. 1943-1952.
- [Rukzio07] Rukzio, E. (2007), 'Physical Mobile Interactions: Mobile Devices as Pervasive Mediators for Interactions with the Real World, Ph.D. Thesis', (Ludwig-Maximilians-Universität München). pp. 11-12.
- [Shirazi09] Shirazi A., Winkler A., Schmidt A. (2009), 'Flashlight interaction: a study on mobile phone interaction techniques with large displays', *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Bonn, Germany: ACM), Article n°93.
- [Speckmann08] Speckmann, B. (2008), 'The Android mobile platform (Master Degree Thesis)', (Eastern Michigan University). pp. 16-38, 89.
- [Vajk08] Vajk T., Coulton P., Bamford W., Edwards R. (2008), 'Using a mobile phone as a "Wii-like" controller for playing games on a large public display', *International Journal of Computer Games Technology* (Hindawi Publishing Corp.), Article n°4.
- [Zhong09] Zhong Y., Li X., Fan M., Shi Y. (2009), 'Doodle space: painting on a public display by cam-phone', *Proceedings of the 2009 workshop on Ambient media computing* (Beijing, China: ACM), pp. 13-20.

# Utilização de Dados Fisiológicos na Avaliação de Aplicações Móveis

Luís Duarte                      Marco de Sá                      Luís Carriço  
 LASIGE e Departamento de Informática, Universidade de Lisboa  
 Edifício C6, Campo Grande,  
 1749-016, Lisboa, Portugal  
 lduarte@lasige.di.fc.ul.pt      {marcosa, lmc}@di.fc.ul.pt

## Sumário

*Os ambientes e aplicações móveis têm sido alvo de pesquisa extensiva, com foco nas questões de usabilidade e experiência de utilização. Métodos como gravação de vídeo, Feiticeiro de Oz ou etnografia são amplamente usados para cobrir as referidas questões. Porém, estes métodos dependem de dados observacionais ou emoções expressas pelos utilizadores em situações pós-teste, que poderão não corresponder à verdade. A utilização de modalidades de interacção fisiológica é um tema que ganha ímpeto no qual sinais biológicos provenientes dos indivíduos são usados como meio de interacção com um sistema ou aplicação. Este tipo de interacção permite aos investigadores e programadores o acesso a dados que estariam ocultos caso utilizassem as técnicas de análise convencionais acima referidas. Este artigo descreve o uso destas técnicas de interacção em ambientes móveis através do uso de um electrocardiograma portátil de forma a analisar o ritmo cardíaco dos utilizadores enquanto interagem neste contexto.*

## Palavras-chave

*Interfaces Fisiológicas, Ambientes Móveis, Avaliação.*

## 1. INTRODUÇÃO

Uma das fases mais importantes durante o processo de desenvolvimento de uma aplicação interactiva é a sua avaliação. Existem variados exemplos de técnicas de análise de usabilidade, tais como etnografia, Feiticeiro de Oz [Maudsley93] ou prototipagem rápida [Isensee66]. Todos estes métodos podem ser aplicados na avaliação de aplicações móveis, apesar de alguns serem mais adequados à análise em laboratório enquanto outros podem ser aplicados nos contextos reais de utilização das aplicações [Sá08]. Porém, estes dependem quase na sua totalidade de dados observacionais, ignorando uma significativa quantidade de informação que é expressa pelos utilizadores e que permanece oculta para estas técnicas. Estes dados consistem em reacções fisiológicas involuntárias que são manifestadas pelos utilizadores através do seu corpo. É, então, necessário considerar a inclusão de mecanismos capazes de lidar com este tipo de informação.

A interacção fisiológica é um subdomínio da interacção homem-máquina não-convencional que foca a utilização de sinais biológicos e estímulos do corpo como alternativa ou complemento a técnicas de interacção tradicionais [Beckhaus04]. Estas modalidades podem fornecer informação determinante para compreender a sua experiência de utilização e problemas de usabilidade em aplicações. Neste artigo apresentamos um conjunto

de módulos que permitem a recolha, processamento e visualização de dados fisiológicos. Estes módulos foram integrados numa framework de prototipagem de aplicações móveis de forma a complementar as técnicas de análise de usabilidade existentes na mesma. Realizamos uma experiência onde tentámos perceber se é possível identificar flutuações nos sinais fisiológicos dos indivíduos ao interagirem com diferentes versões de um mesmo protótipo. Os resultados obtidos indicam que existem alterações nos sinais fisiológicos, o que nos leva a concluir que pode existir uma relação entre as reacções dos utilizadores e a usabilidade dos protótipos utilizados.

## 2. TRABALHO RELACIONADO

A avaliação é uma das fases mais importantes durante o desenvolvimento de uma aplicação [Duh06] [Jones06]. Quando se consideram ambientes móveis, existe um conjunto de factores (e.g. tamanho do dispositivo, mudança de contexto, luminosidade ambiente, barulho ambiente, etc.) que emerge e dificulta o desenho de interfaces, mas que não são aplicáveis em ambientes Desktop. Por exemplo, o uso da técnica Feiticeiro de Oz [Davis07] com protótipos móveis é extremamente complexo. Seria necessário seguir os utilizadores enquanto estes testam os protótipos e alterar os cartões no decorrer da tarefa. No que respeita à prototipagem, existem algumas frameworks e ferramentas destinadas ao efeito, como por exemplo as que são apresentadas em

[Hartman06], ou MobPro [Sá08-2]. Porém, apesar do esforço em contemplar as necessidades e características dos utilizadores bem como as do contexto de utilização, pouca ênfase é dada ao contexto individual dos utilizadores enquanto interagem. Recentemente, emergiram técnicas de interacção alternativas, como o uso de sensores para detectar as características do ambiente ou dos indivíduos através da captura de sinais – as interfaces fisiológicas [Anttonen05].

As modalidades de interacção fisiológica são aplicadas em domínios diversos, como por exemplo as artes, entretenimento ou na medicina [Anttonen05] [Beckhaus04]. Estas interfaces são usadas para obter sinais biológicos dos utilizadores para usufruir desta informação nas aplicações, podendo identificar movimento ou emoções. A maior parte do trabalho existente nesta área está relacionado com o uso dos sinais como meio de adaptação do conteúdo de uma aplicação face ao estado emocional do utilizador. Exemplos desta prática passam pelo feedback através de descargas eléctricas em videojogos [Kruijff06] ao ajuste automático da dificuldade de um jogo baseado no nível de stress do jogador. Existe também um trabalho sobre a utilização da condutividade da pele em conjunto com a pressão aplicada aos botões de forma a adaptar a dificuldade de um jogo [Mandryk06]. Os resultados mostram que é possível detectar diferentes estados emocionais com precisão.

### 3. CONTEXTUALIZAÇÃO - MOBPRO

MobPro é uma framework de prototipagem e avaliação especialmente direccionada ao desenho de interfaces com o utilizador para ambientes móveis [Sá08-2]. É composta por um conjunto de ferramentas que se complementam e oferecem aos programadores os meios necessários para criarem, ajustarem e testarem protótipos de interfaces fora do laboratório. Mais informação sobre a framework pode ser consultada em [Sá08].

#### 3.1 Análise e Recolha de Dados

O MobPro inclui um conjunto de mecanismos para contemplar a recolha de dados de interacção e avaliação dos protótipos. De forma a dar suporte a este processo, disponibilizamos as seguintes funcionalidades:

- (1) Questionários de usabilidade que podem ser mostrados ao utilizador em eventos específicos.
- (2) Anotações e registos de diário que podem ser adicionados em cada ecrã ou componente.
- (3) Mecanismo de logging que permite a recolha automática e cronometrada de dados de interacção à medida que o utilizador manipula o protótipo.

De forma a dar suporte à análise dos dados recolhidos o MobPro inclui duas abordagens distintas. A primeira permite aos desenhadores navegarem e reverem as respostas dos utilizadores aos questionários, as anotações e os registos de diário. A segunda abordagem de análise inclui um mecanismo que funciona como um reproduzidor

de logs, oferecendo as funcionalidades necessárias para rever e analisar os logs através de um módulo que permite a emulação das interacções do utilizador com o protótipo.

### 4. RECOLHA DE DADOS FISIOLÓGICOS

A integração da recolha e manipulação de dados fisiológicos na framework MobPro levou à concepção de um conjunto de módulos específicos para o efeito. Estes devem cobrir aspectos distintos, tais como disponibilizar controlos adequados à visualização dos dados, estruturas de dados flexíveis para o armazenamento dos mesmos, entre outros. Os módulos que foram identificados e idealizados são:

- Aquisição de Dados: responsável pela recolha e interpretação de dados fisiológicos dos dispositivos;
- Representação de Dados: responsável pela definição de estruturas de dados usadas para manipular e guardar dados fisiológicos em tempo de execução, e do seu armazenamento em ficheiros ou bases de dados;
- Interface com o Utilizador: disponibiliza os controlos necessários aos analistas para poderem visualizar os dados fisiológicos recolhidos.

Com os módulos anteriormente descritos criámos uma ferramenta de análise fisiológica para ambientes Windows Desktop e Windows Mobile que permite a visualização de dados fisiológicos em conjunto com logs de interacção. Esta ferramenta permite aos utilizadores observar gráficos de dados fisiológicos bem como obter dados estatísticos simples como a média ou o desvio padrão para qualquer modalidade fisiológica. Os logs fisiológicos armazenam o instante inicial e final de uma experiência bem como estampilhas temporais de modo a possibilitar a navegação pelos dados quando são analisados. Uma das vantagens desta ferramenta é a possibilidade de visualizar informação que foi anteriormente gravada ou monitorizar o utilizador ao mesmo tempo que este vai interagindo com os protótipos. Esta característica é particularmente útil para avaliação in-situ na qual os analistas podem detectar problemas e ajustar os protótipos no local, baseando-se nos dados observados.

### 5. EXPERIÊNCIA

Conduzimos uma experiência com o intuito de validar os módulos de dados fisiológicos e a abordagem escolhida na framework MobPro.

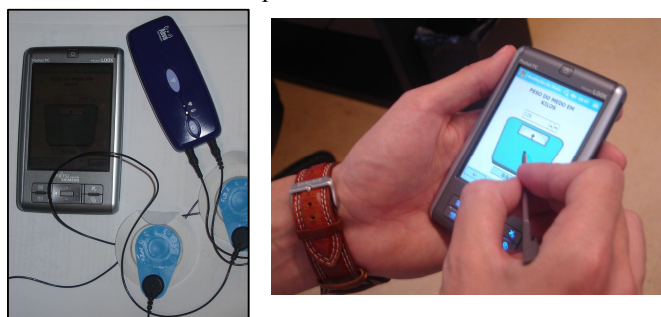
#### 5.1 Abordagem Experimental

De forma a perceber se conseguiríamos detectar algum problema de usabilidade baseando-nos nos logs de interacção e nos dados fisiológicos, primeiro foi necessário identificar as alterações no ritmo cardíaco dos utilizadores e mapear essas alterações nos eventos obtidos dos logs. De forma a facilitar este processo, usámos duas versões de um protótipo da mesma aplicação móvel.



A aplicação móvel foi criada no âmbito de um projecto de psicoterapia para o tratamento de patologias relacionadas com o medo. A primeira versão da aplicação continha três ecrãs interactivos e botões de navegação, nos quais os utilizadores podiam quantificar a intensidade dos seus medos, usando aproximações diferentes:

- No primeiro ecrã, a imagem de um edifício pode ser esticada, do nível 1 até ao nível 10, através do arrastamento de uma pequena área identificada pela expressão “Aqui!”.
- No segundo ecrã podemos observar uma balança com um contador que pode ser aumentado até 1000 kgs de forma a indicar o “peso” do medo.
- No terceiro ecrã o utilizador pode quantificar a força do seu medo clicando directamente na barra que se encontra do lado esquerdo.



**Figura 1 – Equipamento: Pocket PC, ECG e Protótipo.**

A segunda versão da aplicação é composta pelos mesmos três ecrãs. Porém, neste caso, problemas de usabilidade específicos foram induzidos: redução da precisão de selecção dos valores nos diferentes ecrãs; redução das áreas interactivas em cada ecrã; indução de comportamento aleatório nos botões de navegação.

### 5.1.1 Equipamento & Ferramentas

O equipamento fornecido aos utilizadores consistia num Pocket PC, os protótipos da aplicação descrita e o dispositivo de recolha de dados fisiológicos. Este último foi preparado com eléctrodos e gel condutor de forma a obtermos o melhor sinal possível (Figura 1).

### 5.1.2 Participantes & Procedimento

Todas as sessões de teste foram realizadas no campus de uma universidade. Cada sessão demorou cerca de 10 minutos. Um total de 15 sujeitos participaram nestas sessões. Os participantes tinham idades compreendidas entre os 19 e os 36 anos. Cada teste consistia em duas tarefas. O procedimento adoptado foi o seguinte:

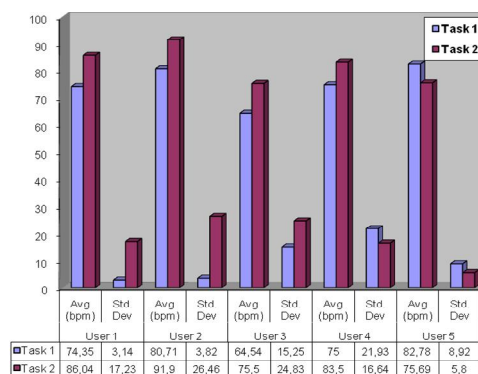
- Pré-Tarefa: os participantes interagiram com uma ferramenta de calendário e navegaram por um album de fotografias presente no Pocket PC.
- Tarefa 1: os participantes usaram o protótipo e foram requisitados para colocar o “valor” do medo a 3

em todos os ecrãs. De seguida, foi-lhes pedido que modificassem todos os valores para 5.

- Tarefa 2: a segunda tarefa é semelhante à anterior, mas desta vez os participantes usam a segunda versão do protótipo, com os vários problemas de usabilidade.

O teste foi dividido em três cenários, com cada participante a ser aleatoriamente escolhido para realizar apenas um deles:

- Cenário 1: os participantes realizaram as tarefas anteriormente descritas enquanto estavam sentados numa sala de aula convencional.
- Cenário 2: os participantes interagiram com os protótipos enquanto caminhavam pelos corredores de um dos edifícios da universidade.
- Cenário 3: os participantes realizaram as tarefas em ambiente livre, caminhando pelo campus ou sentados.

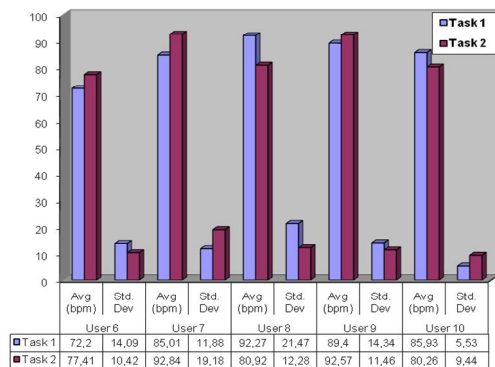


**Figura 2 – Resultados Cenário 1.**

## 5.2 Resultados

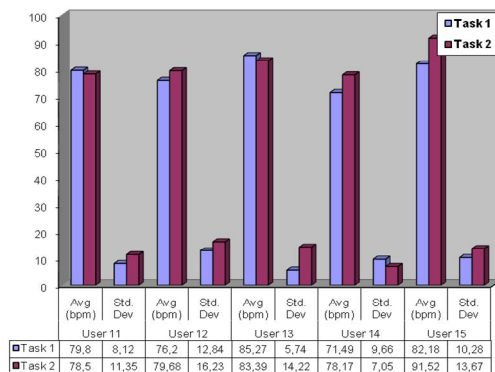
Nesta experiência focámo-nos na identificação e estudo do ritmo cardíaco médio para todos os participantes durante as tarefas e do desvio padrão obtido. O ritmo cardíaco médio permite identificar se uma tarefa afectou mais o utilizador que outra. O desvio padrão, juntamente com a observação cuidada dos logs de interacção e dos logs dos dados fisiológicos permitem identificar se o utilizador foi mais ou menos afectado pelos problemas de usabilidade e os segmentos em que tal aconteceu.

Os resultados para os diferentes cenários podem ser observados nas Figuras 2, 3 e 4. Como consta nas figuras, na segunda tarefa (a coluna à direita de cada par) houve um aumento do ritmo cardíaco médio da maioria dos participantes. Através dos logs de interacção e dados fisiológicos foram identificados os períodos em que se registou uma maior subida do ritmo cardíaco. Estes períodos corresponderam aos instantes em que os utilizadores interagiram com o primeiro e o terceiro ecrã da segunda versão do protótipo.



**Figura 3 – Resultados Cenário 2.**

No caso dos cenários 2 e 3 podemos concluir que o ritmo cardíaco médio aumentou da tarefa 1 para a tarefa 2, mas mais ligeiramente comparativamente ao que se tinha verificado no primeiro cenário. Isto deve-se ao facto de os utilizadores se encontrarem em movimento, o que leva o ritmo cardíaco a aumentar, tornando mais difícil a identificação de flutuações atribuídas a problemas de usabilidade. No entanto, através da inspecção dos logs de interacção e dos dados fisiológicos, foi possível verificar que existe um aumento acentuado do ritmo cardíaco quando os utilizadores interagiram com os ecrãs 1 e 3 do segundo protótipo.



**Figura 4 – Resultados Cenário 3.**

Concluindo, as experiências apresentadas indicam que a utilização de dados fisiológicos em ambientes móveis é uma abordagem válida no que respeita a estudos de análise de usabilidade. Desta forma, concluímos que o uso de dados fisiológicos é mais adequado como complemento a técnicas de análise existentes, fortalecendo assim o processo de avaliação.

## 6. CONCLUSÕES E TRABALHO FUTURO

Este artigo descreveu o uso de modalidades de interacção fisiológica como uma fonte de dados complementar para avaliação de usabilidade em aplicações móveis. O trabalho tem como base uma framework de prototipagem para dispositivos móveis, oferecendo as extensões necessárias de forma a contemplar as novas modalidades de interacção. A necessidade de possuir informação fisiológica a complementar dados de interacção convencionais deve-se ao facto de cada indivíduo

manifestar e expressar informação valiosa nos seus sinais biológicos que não é detectada através de meios de análise convencionais. Foi conduzida uma experiência de forma a testar a nossa hipótese sobre a possibilidade de correlacionar problemas de usabilidade com alterações nos sinais biológicos dos indivíduos. Os resultados obtidos mostraram que a maioria dos utilizadores teve um aumento no seu ritmo cardíaco ao interagir com a versão da aplicação que continha os mecanismos frustrantes.

Como trabalho futuro, planeamos incluir mais modalidades de interacção fisiológica, como por exemplo o nível de condutividade da pele.

## 7. REFERÊNCIAS

- [Anttonen05] Anttonen et al. Emotions and heart rate while sitting on a chair. In Procs of the SIGCHI conference on Human Factors in Computing Systems, pp. 491-499, 2005.
- [Beckhaus04] Beckhaus et al. Unconventional Human Computer Interfaces. In ACM SIGGRAPH 2004 Course Notes, Article no. 18, USA, 2004.
- [Davis07] Davis et al. SketchWizard: Wizard of oz Prototyping of Pen-Based User Interfaces. In Procs of UIST'07, pp. 119-128, 2007.
- [Duh06] Duh et al. Usability Evaluation for Mobile Device: A Comparison of Laboratory and Field Tests. In Mobile HCI'06. 2006.
- [Hartman06] Hartman et al. Wizard of Oz Sketch Animation for Experience Prototyping. UbiComp 2006, 17-21, 2006.
- [Isensee66] Isensee et al. The Art of Rapid Prototyping Intl. Thomson Computer Press, London.
- [Jones06] Jones et al. Mobile Interaction Design. 2006, West Sussex, England: John Wiley & Sons Ltd.
- [Kruijff06] Kruijff et al. Using neuromuscular electrical stimulation for pseudo-haptic feedback". In Procs of the ACM Symposium on Virtual Reality Software and Technology, pp. 316-319, Cyprus, 2006.
- [Mandryk06] Mandryk et al. A continuous and objective evaluation of emotional experience with interactive play environments". In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, ISBN: 1-59593-372-7, pp. 1027-1036, Canada, 2006.
- [Maudsley93] Maudsley et al. Prototyping an intelligent agent through Wizard of Oz. Interchi '93 Conference Procs, pp 277-284, 1993.
- [Sá08] Sá et al. Lessons from Early Stages Design of Mobile Applications. In Procs of MobileHCI'08, pp. 127-136, 2008.
- [Sá08-2] Sá et al. A Mixed-Fidelity Prototyping Tool for Mobile Devices. In Procs of AVI'08, 225-232, 2008.

# Automatic Volume Adjustments for Mobile Media Players

Tiago Reis

Luís Carriço

Carlos Duarte

LASIGE e Departamento de Informática, Universidade de Lisboa  
Edifício C6, Campo Grande, 1749-016, Lisboa, Portugal  
[treis@lasige.di.fc.ul.pt](mailto:treis@lasige.di.fc.ul.pt) [lmc@di.fc.ul.pt](mailto:lmc@di.fc.ul.pt) [cad@di.fc.ul.pt](mailto:cad@di.fc.ul.pt)

---

## Abstract

*The automatic adjustment of mobile user interfaces to the contexts in which they are used presents a possible solution for the interaction difficulties introduced by the constantly mutating characteristics of such contexts. Here we present a prototype of a portable media player. This player automatically adjusts its volume according to the users' preferences and the noise of the surrounding environment. We focus on the details of the volume adjustment algorithm and its – laboratorial and contextual - user centered evaluations. We detail the problems found on each phase of the evaluation procedure, and present a set of Wizard of Oz experiments exploring possible solutions for those problems.*

## Keywords

*Mobile HCI, Automatic Volume Adjustments, Context Awareness, User Centered Design, Contextual Evaluation.*

---

## 1. INTRODUCTION

Nowadays, mobile devices are strongly integrated in peoples' lives. The ubiquitous nature of these devices enables humans to use them in an enormous variety of contexts, defined by various sets of dynamic characteristics (contextual variables) that heavily affect user interaction. However, the differences amongst users and the frequent context mutations, which occur during device utilization, often result on users' adaptation to both contexts and interfaces available, and not the other way around. Most mobile user interfaces are unable to adapt effectively and automatically to the mutations of their utilization contexts, introducing interaction difficulties [Barnard07, Reis08, Schmidt99]. Accordingly, it is necessary to explore new approaches for user interface design, aiming at usability improvements. To achieve this, applications must be constantly aware of the relevant aspects of their utilization contexts and respective mutations, naturally providing their users with adequate interaction modalities, combining and configuring them accordingly.

Audio-based mobile interaction processes are significantly affected by noise, a contextual aspect that varies considerably on mobile scenarios. Amongst these processes, the use of portable media players is a reality that remounts to the late 70's when the Sony Walkman became available to the public. Nowadays, these players evolved and became an integrated component of several mobile devices, which serve other purposes as well. The volume controller on portable media players also suffered several changes since the late 70's volume wheel. This component evolved to physical buttons and, later, to virtual controls presented on touch screens. The substantial utilization of touch screens eradicated the tactile awareness and feedback provided by mechanical controls, making it often impossible to control such devices without having

to look at them. Nevertheless, some headphones include small controllers (e.g.: volume, forward, etc.) placed on their wire in order to reduce the users' need to look at the device in order to interact. Despite the evolution of portable media players and respective controls, noise still affects interaction considerably, and the mutations of this context variable often result on manual volume modifications. The reduction or eradication of these modifications may introduce significant improvements to the portable media players' user experience.

From simple headphones and headsets with physical noise canceling [Review10] to headphones and headsets that include very complex algorithms of noise elimination [JawBone10, TheBoom10], different solutions have been proposed and created in order to reduce the impact of noise on audio interaction. Noise canceling and elimination are paramount in critical scenarios, such as: communicating on a landing strip. However, this solution demands an aggressive noise elimination that is extremely complex and, consequently, extremely expensive, sometimes making the headphones themselves more pricey than regular mobile devices [Review10, Jawbone10, TheBoom10]. Moreover, noise canceling solutions may become dangerous in every-day mobile contexts. For instance, while a user is walking on the street, if (s)he is completely inhibited from hearing the cars passing by, (s)he can be injured. Accordingly, noise-based automatic volume adjustments present a more affordable and safe solution for contexts in which the environmental noise varies significantly enough to force users to manually modify the volume of their portable media players. Moreover, several mobile phones function also as portable media players, and all of them include a microphone that can be continuously used in order to monitor the environmental noise.

The work presented in this paper addresses the contextual adaptation of volume on portable media players. It is materialized on a functional prototype and a set of user-centered studies. The prototype is a noise-aware media player that attempts to reduce the manual volume modifications generated by noise, gathering noise levels from the surrounding environment, and automatically adjusting the volume accordingly. These adjustments also consider the users' preferences, which are transparently identified whenever a manual volume adjustment is performed. The studies conducted focused on the details of the automatic volume adjustment algorithm and its employment on real life scenarios. These studies revealed a strong user acceptance of both the concept proposed and its implementation. However, for specific situations where the users engage in conversations, our implementation of the concept generated slightly uncomfortable reactions on the users. Building on top of such results, we also present a set of Wizard of Oz studies, conducted in order to evaluate a set of possible solutions for this problem.

## 2. THE PROTOTYPE

The prototype created includes a noise monitor and a continuous preference-based volume adjustment algorithm. Logging mechanisms were also implemented in order to ease the contextual evaluation. The prototype is an application for devices running Windows Mobile 5. The noise monitor employed is the one available in [Mitchell07], which gathers sample values from the device's microphone and calculates loudness values in a spectrum that varies from 0 to 127.5 root-mean-squares (RMS). The algorithm responsible for the automatic volume adjustments considers noise as the primary context variable. It was specifically designed for media control scenarios in which the users are wearing headphones. Accordingly, the sound produced by the media being played does not influence the noise monitoring process, enabling the continuous monitoring of the environmental noise. The automatic volume adjustments are applied continuously within a volume spectrum that may vary from 0 to 100. The algorithm is configurable, 4 parameters are accessible to the users (Figure 1): minimum, maximum, sensibility, and volume step. The first two define the minimum and maximum volumes that can be set by the algorithm. These boundaries are defined in order to avoid adaptations that are inconvenient for the users (e.g.: setting the volume to low due to the absence of noise). The sensibility parameter ranges from 1 to 5, defining **when** the algorithm reacts. This value is employed on the definition of the coefficient dividing the noise spectrum into a set of discrete noise levels. Each of these levels defines a line on the adaptation table used by the algorithm and has a correspondent volume level that is defined by the volume step. The volume step parameter ranges from 1 to 10. This value defines **how** the algorithm reacts, corresponding to the amount that the volume is automatically increased/decreased, whenever the noise of the surrounding environment goes up/down one level. Accordingly, the automatic volume adjustment algorithm behaves ac-

cording to a dynamic table of pairs noise/volume. The number of lines on this table is defined by sensibility, and the correspondent volume levels are defined by the volume step, varying between the boundaries defined by the minimum and maximum values.

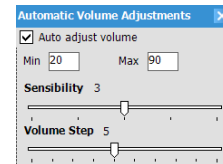


Figure 1. Configuration screen of the algorithm.

This algorithm is also preference-based. The user's preferences are indirectly expressed whenever (s)he performs a manual volume modification. When this happens, the algorithm assumes that the volume set represents the user's preference for the noise captured at that moment, overriding his/her previous preference, and instantaneously modifying the table noise/volume. From that moment on, the adaptation continues according to the modified table, which is saved on a XML file for posterior use.

The logging mechanisms register the users' and algorithm behaviors on XML files. Every action is registered and associated with a contextual stamp, which includes time and noise-related information.

## 3. EVALUATION

The evaluation process comprised two strongly user centered procedures (laboratorial and contextual). The main goals of these procedures were to understand: (1) if the automatic volume adjustments reduce the number of manual volume adjustments performed in order to mitigate noise disturbances; (2) if the configuration parameters were adequate; and, (3) if the users believed the concept proposed to be useful.

### 3.1 Laboratorial Evaluation

The laboratorial evaluation involved four users: 3 male, 1 female, with ages between 20 and 29 years old. The experiment conducted in order to fulfill the first goal comprised two phases. On the beginning of both phases the users were asked to set the volume to a comfortable level. On the first phase, the users listened to a song while an audio stream was played in order to simulate the background noise. This stream contained sounds that are typically found on real mobile interaction scenarios (e.g.: birds, cars driving by, trains, noise generated by people, etc.). During this phase the users performed in average 5 manual volume adjustments in order to mitigate the noise disturbances. On the second phase, the automatic volume adjustments were activated with the default values, significantly reducing the number of manual volume adjustments. The users listened to the same song while the same audio stream was played, averaging less than 1 manual volume adjustment. The manual volume adjustments performed on this phase emerged strictly because 2 of the

users were not comfortable with the default maximum value of the algorithm.

Following, in order to fulfill the second goal, the users were presented to the configuration screen of the automatic volume adjustment. All users considered the minimum and maximum volume parameters of utmost importance, emphasizing that they allow an important level of personalization. Regarding the sensibility and volume step, users were confused with the existence of these two parameters. All of them were expecting just one parameter. After experimenting different configurations for these two parameters, all the users reported that high sensibility values should correspond to low volume steps, and low sensibility values should correspond to higher volume steps. The users reported to be very comfortable with the default values for the sensibility and volume step, considering them the most adequate combination of all the ones they have experimented. Moreover, they stated that this feature should either be available through one control only, or not be available at all.

Finally, when asked about the usefulness of the automatic volume adjustments all the users considered them very useful, and were very enthusiastic about the possibility of integrating our prototype in their every-day lives.

### 3.2 Contextual Evaluation

The evaluation procedure conducted in real life scenarios involved six users: 3 male, 3 female, with ages between 18 and 35 years old. The users involved included our prototype on their day-to-day lives during a two week period summing a total of 40 utilizations on real contexts. In the end of the process, the users participated on an informal interview where they provided their feedback and opinions about the automatic volume adjustments, emphasizing situations where these were, and were not, satisfactory.

All the users reported that the number of manual volume adjustments performed in order to mitigate noise disturbances was significantly reduced when compared to their own portable media players. Users reported several situations where they had to move between contexts with significantly different environmental noise values, without the need to manually modify the volume configuration. Such situations include going from home to the users' workplaces, having to walk on the street, ride public transportations, get in and out of different buildings, etc. The log analysis corroborates these reports, showing no manual volume adjustments during long periods of time (until 2 hours), characterized by very discrepant environmental noise values.

The users reported other contextual aspects that generate manual volume adjustments. Amongst these, they emphasized their emotional states, the song being played, and the primary task they were involved in (e.g.: reading a book) to affect their preferences regarding the volume set. Nevertheless, all the users stated that these situations were well covered by the prototype behavior, namely by the preference base of the algorithm.

There was one particular type of situation where users felt uncomfortable with the algorithm's behavior. Whenever the users were interrupted by people, engaging in conversation, the volume would increase due to the increasing environmental noise generated by the conversation. These situations often culminated on a manual volume modification or on users removing their headphones. The users did not consider this to be a major problem. They all reported that the improvements achieved on the majority of the scenarios justified such inconvenience. These uncomfortable situations are also patent on the logs gathered, where in some situations of increasing environmental noise the users manually set the volume to mute.

Despite being able to modify all the algorithms' parameters, the users only personalized the maximum and minimum volumes. Again, they reported some confusion regarding the existence of the two remaining parameters. Nevertheless, all the users reported smooth and accurate automatic volume adjustments, showing a high satisfaction regarding the default values of such parameters. The logs gathered corroborate these statements, showing modifications only for the maximum and minimum parameters.

Regarding the usefulness of the automatic volume adjustments, all the users acknowledged this functionality to be very useful. All of them reported improvements of the interaction process, namely a significant reduction on the number of interactions with the device, and, consequently, a better user experience. Moreover, when the users started using their own portable media players again they stated that once familiarized with the automatic volume adjustments having to go back to performing them manually became significantly more "annoying".

### 4. WIZARD OF OZ EXPERIMENTS

Despite the strong user acceptance of our prototype, its contextual evaluation procedure revealed uncomfortable situations for the users. These situations were generated whenever the users engaged in conversations with other people. In order to explore automatic solutions for this problem, a series of Wizard of Oz experiments was conducted. During the process, both the volume and the automatic volume adjustment algorithm were remotely controlled in order to simulate the particularities of each experiment.

The first experiment considered the perfect identification of the beginning and end of conversations involving the user. The experiment focused the behavior the prototype should assume during the conversations the users engage in. Three alternatives were experimented: 1) pause; 2) muting the volume; and, 3) lowering the volume to the minimum value, and disabling the volume adjustment algorithm. Four users preferred the first solution, one preferred the second, and one voted for the third, suggesting that such behavior should be configurable since its selection is strongly affected by the users' preferences. Regardless of the behavior assumed by the prototype, the perfect identification of the beginning and end



of conversations involving the user revealed a strong user acceptance; all the users reported this to be the perfect solution for the problem.

However, the fully reliable identification of such situations is unfeasible if we only consider the technology available in most mobile devices nowadays (one microphone and limited processing capabilities). Accordingly, the following two sets of experiments considered some problems that could emerge from the employment of solutions based on voice activity detection, namely the existence of: 1) false positives (e.g.: identifying a conversation that involves the user when the user is not involved in the conversation); and, 2) false negatives (e.g.: not identifying a conversation involving the user when the user is actually involved in a conversation).

The experiments considering only the existence of false positives on the identification of conversations where the users participate showed that unless the mechanism responsible for this identification is fully (100%) reliable the uncomfortable situations increase when compared to the actual implementation of our prototype. Accordingly, the existence of false positives was strongly rejected by all the users. The experiments considering only the existence of false negatives, presented improvements when compared to the actual implementation of the prototype. The users reported that the identification of some conversations and the consequent automatic behavior improved the interaction process, reducing the manual volume adjustments. Such reduction is directly proportional to the reliability of the mechanism responsible for the identification of conversations where the users participate. However, the existence of false negatives leaves us with the same problem found during the contextual evaluation, implying the creation of a mechanism to overcome the interaction difficulties they introduce.

Accordingly, we conducted a new set of experiments focusing the manual control of the situations where the users engage in conversations. Different solutions were experimented for the activation/deactivation of the behavior assumed by the prototype during a conversation, namely: 1) the use of a button available on the wire of the headphones; 2) the use of a button on the device; 3) the use of gestures on the device's screen; and, 4) the use of voice commands. After trying all the solutions proposed, users rated them according to their usability and compared them to the solutions adopted during the contextual evaluation. The use of voice commands was considered very poor regarding its usability. The users reported to feel very uncomfortable talking to the device, especially in public places where there are more people. The use of gestures on the device screen and the use of a button on the device were considered poor regarding their usability. The users stated that solutions considering the direct manipulation of the device, which is often on their pockets or bags, are not adequate. The control through a button available on the headphones' wire was considered the most usable solution. When asked to compare the solutions proposed with the ones adopted during the contextual

evaluation, all the users still considered the use of a button available on the headphones' wire as the most adequate solution. However, they all stated that the second most adequate solution would be the manual control of the volume through a volume controller placed on the headphones' wire. Finally, the users stated that removing the headphones is a better solution than voice commands.

Overall, the above presented experiments enabled us to conclude that a fully reliable solution for the identification of situations where the users engage conversations would solve the problem found during the contextual evaluation procedure. However, this is currently unfeasible from a technological point of view. Accordingly, the manual control of such situations presents the most adequate solution for the problem.

## 5. CONCLUSION AND FUTURE WORK

In this paper we described a media player that automatically adjusts its volume according to the environmental noise and to the users' preferences. The prototype presented was evaluated through two strongly user centered procedures. These procedures revealed a strong user acceptance of the prototype, pointing some issues that should be improved. Following, a set of Wizard of Oz experiments was conducted in order to explore a set of possible solutions for these issues. Our future work comprises the development of plug-ins for mobile media players. These shall enable automatic noise- and preference-based volume adjustments, employing the knowledge assembled on the experiments here presented.

## 6. ACKNOWLEDGEMENTS

LaSIGE and FCT. Multiannual Funding Programme and individual scholarships SFRH/BD/44433/2008.

## 7. REFERENCES

- [Barnard07] Barnard, L., Yi, J S., Jacko, J., Sears, A.: Capturing the effects of context on human performance in mobile computing systems. In *Personal and Ubiquitous Computing*, Vol. 11, 2, pp. 81 - 96.
- [Reis08] Reis, T., Sá, M., Carriço, L.: Multimodal Interaction: Real Context Studies on Mobile Digital Artefacts. In *Procs. of the 3<sup>rd</sup> HAID*.
- [Schmidt99] Schmidt, A., Aidoo, K.A., Takaluoma, A., Tuomela, U., van-Laerhoven, K., and Van-de-Velde, W.: Advanced interaction in context. In *Procs. of the 1<sup>st</sup> HUC*, pp. 89-101.
- [Review10] Review on noise canceling headphones. Available at: [http://www.seatguru.com/articles/noise-canceling\\_review.php](http://www.seatguru.com/articles/noise-canceling_review.php)
- [JawBone10] The Jawbone Headset. URL: <http://eu.jawbone.com/epages/Jawbone.sf>
- [TheBoom10] Boom <http://www.theboom.com/>
- [Mitchell07] Mitchell, C.: Mobile Apps: Adjust Your Ring Volume For Ambient Noise. In *MSDN Magazine*.

## **Sessão III**

### **Concepção e Desenvolvimento da Interacção**

Aveiro, 13, 14 e 15 de Outubro de 2010



# The Genio Centred on the Human Being

## A Methodology for Fast Development Centred on the Human Being

Duarte Teixeira

IST/UTL

Lisboa

duarte.teixeira@gmail.com

Mário Rui Gomes

DEI/IST/UTL

Lisboa

mario.gomes@tagus.ist.utl.pt

### Abstract

*The development of an Information System only makes sense if it achieves its main purpose: solving the problem or set of specific problems for which it was designed and conceived. However, it is increasingly imperative that the development has to be made due to concrete specifications, which results on the solution to a specific problem, and not on a solution to a type of problem.*

*This paper presents in detail the GenioCSH, a methodology for development of Information Systems focused on the Human Being that, with the aid of a RAD (Rapid Application Development) tool, allows improvements in the development process based on concrete specifications.*

*The practical application of the GenioCSH methodology in the development of two real projects resulted in a reduction on development time and considerably on the volume of incidents and requests for clarifications made in comparison with the estimated values. Information Systems developed with the application of the GenioCSH methodology proved that besides being able to fill existing needs, did not lead to increased costs in their learning, thus responding in a positive way, and since the first day, the challenges for which were developed.*

### Keywords

*Usability, Users, Interaction, Development, RAD, PDCA.*

## 1. INTRODUCTION

In the beginning of the “Computer Era” the use of computers and “high-tech” electronic equipments required consistent skills from their users, sometimes only in the reach of the most educated. The Human Being was the one adapting to the systems (or machines), the one having to changes his habits and ways of thinking, towards the way the system worked.

With the evolution of times, many progresses where made to make the systems less harsh and more usable. Their omnipresence and diversity took the initial adaptation upside-down and the word “Usability” started making part of the design and development process.

In our days, Usability represents a huge factor in the decision making process. Features, price and thrust factors related with the supplier will forever have their importance, but what users and managers will always think and question themselves about is: “*what’s a system full of features worth for if I can’t use them?*”

The work stated here present a methodology that can be used on the development of IS (Information Systems) centred on the human being, with the support of a RAD

(Rapid Application Development) tool. Keeping the fast development, the methodology improves the quality and robustness of the final solution, increases the usability and functionality of the system as well as helping providing the customer with the specific solution for his specific need and wish.

This methodology brings the users and all other actors actively into the development process, each actor for his specific roll, inviting them to take part in the development of “*The Solution*” to the problem, instead of “one solution” to the problem.

## 2. RELATED WORK

There are several ways to make sure that a certain IS will reach the desired levels of usability, nevertheless, each of them have advantages and disadvantages that have to be considered and deliberate due to the context [1] [2].

On an early stage a research took place to identify significant related work on this scientific field. Due to the research, several methods and models were studied and analysed, but only a few match the common goals with the work here presented.

## 2.1 RITE - Rapid Iterative Testing and Evaluation

A methodology in some parts similar to the traditional method, differing in the fact that it executes individual tests with users, followed by changes to the presented solution. These changes will be tested in the next tests, but, their execution can take too long to be complete or even don't be able to be executed in time [3] [4] [5].

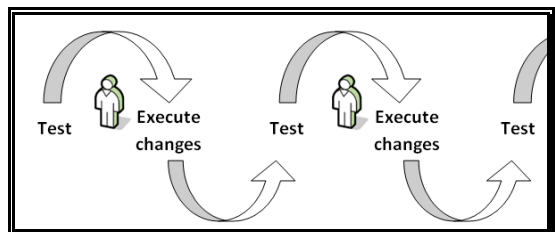


Figure 1 - The RITE Method

With RITE, the tests made with users are made in sequence so that the proposals made by one user can be part of the next user solution.

### ADVANTAGES

The fact that it will be, iteratively, made several changes to the solution initially presented, it makes the solution constantly to be refined and improved, reaching an acceptable state.

Apart from executing several changes through the user testing "session", these changes can automatically be tested and analysed in the following tests, so in the end of the user tests, the solution initially presented gathered several changes, changes that have already been tested and analysed by the users.

### DISADVANTAGES

The approach can only be applied on complete functional solutions.

The solution presented has to be in a late state of development (and be functional), reducing in that way, the space for deep changes to take place on the actual solution.

The changes proposed by the users can take too long to be executed, delaying or even making it impossible to continue with the tests.

The proposals made by the users can divert the actual solution from the objective initially proposed, bringing the solution to an undesirable condition.

## 2.2 OVID - Object, View and Interaction Design

This method reinforces the idea that the visual aspects and the interaction with the user are important, but not the most important aspect of the interface.

The OVID method says that objects, their relations, their behaviours and how these support the execution of the

necessary tasks by the user are the most important aspects of the interface and of the design model.

The method is composed by three models: the user model, the designer model and the programmer model. The user model consists on the interpretation of the user on what he thinks it will be (or should be) the final solution. The designer model consists on the existing idea of what the users should see and be able to interact with when they'll use the developed solution. At last, the programmer model consists on what will be developed and the way it will be developed.

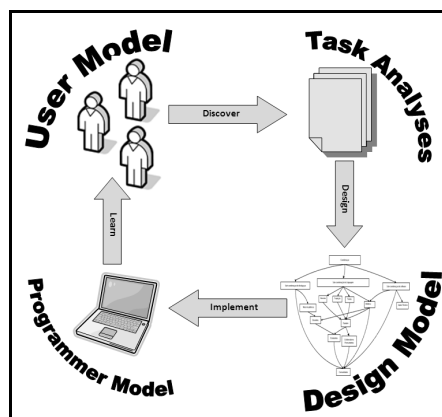


Figure 2 - The OVID Method

### ADVANTAGES

The developed solution is a reflection of the user perspective since he's the one giving the necessary input so that the solution starts to be idealized, designed and developed.

The (user) specification is presented in a simple way, due to the fact that the user has no knowledge in the technical/programming/developing area. These specifications can be used by the programmers to directly structure the solution classes, since the user specification already presents all the possible relations and interactions.

The notations used by the user aren't limited to any specific format, giving the development team the capacity to adapt it to any notation or tool used in the development.

### DISADVANTAGES

As seen above, the method allows that any format of specification can be made by the user, and that it will be the source and initial point to the conception of all other models involved. This also means that, a necessary change made to any model represents the need to execute changes in all models.

## 2.3 MUSE - Method for Usability Engineering

This method is mostly designed for developing single user systems [6].

With MUSE, all systems that the user interact with must be analysed and studied so that similarities and common

patterns be identified and be included on the system to be developed.

All the systems similar to the one being designed should be analysed, but not only these systems matter in this method. Even other systems or other tools (computerised or not) that the user has contact with on his daily profession should be analysed.

The MUSE method defends that developing systems with similarities to those other systems familiar to the user, helps the learning process and the recognition of the system functions, their location and their objective [6] [7].

#### ADVANTAGES

Using the MUSE method on the development of IS origins an assimilation of the processes, on the task done by the user and on the way both should be done. Due to the studies done to other systems present on the daily activity of the user, the system developed with this method present similarity between the new and the existing systems, resulting on a raise of familiarity, productivity and satisfaction, and consequently on a reduction on the need for training, on the technical assistance and on the user errors.

#### DISADVANTAGES

The value of applying this method to the development of a IS reduces as the number of systems user departs from 1. The fact of existing more than 1 user for the system turns the study and analyses phase larger than it's supposed to be, but even more undesirable, it results on a very impersonal and very vague model with generalised design instead of matching the user.

## 2.4 WISDOM - Whitewater Interactive System

### Development with Object Models

This method represents an evolutionary methodology centred on the user, with fast prototyping and turned into the development made by small teams, all in charge of doing all the rolls in the development process. The development team elements should be able to make the design, development and maintenance of the system [8].

#### ADVANTAGES

The WISDOM method is planned for the development of interactive systems by small companies or small development teams, since it allows them to be flexible, provide quick response, improve and use their creativity directly on the development process.

#### DISADVANTAGES

Despite the WISDOM method supports the communication between the development team, the use of proposed notation, it can be considered very complex and specific.

## 3. THE GENIOCSH METHOD

The methodology here described emphasis the need for developing solutions that fit the needs of the client, solutions that represent an available tool to solve the customer

problem, and solve it the fastest, the easiest, the more effective and more profitable way.

GenioCSH stands for “*Genio Centrado no Ser Humano*” (in Portuguese) meaning “*Genio Centred in the Human Being*”. The name Genio comes from the RAD tool used on the design of the methodology, a proprietary tool owned by Quidgest<sup>1</sup>. As said above, the design and tests done with this methodology (and presented here), where done with the mentioned RAD tool.

Main goals of GenioCSH are too keep the fast development of IS, doing it now centred on the Human Being, developing the perfect and specific solutions for a specific problem/client, and not one solution to a kind of problem.

To attain all wished results with the use of GenioCSH it's necessary that all actors participate in the design and development process. Users, development team, project managers and supervisors, all have their roll in the methodology.

With GenioCSH a group of individual test should be made, called as “*Block Tests*”, where some of the IS features are tested and where the users are invited to propose changes to the tested features and to the system in general. Then these changes are classified, first by the user and secondly by the development team. After the classification some changes are immediately put under development so they can be finished as soon as possible, so they can be introduced into the system tested by another user of the same block test. This way, proposed changes can immediately be tested by the following users of the block test.

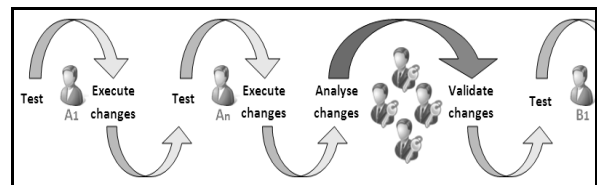


Figure 3 - GenioCSH Method

At the end of the block test (after all individual tests are complete), the proposed changes are submitted to a validation process, done by the key users, project managers and/or supervisors. In the validation the proposed changes should be accepted or rejected. Only with the validation it's possible to check if a change or group of changes were diverting the actual solution from the desired one, going in an opposite direction of what's desired by the customer.

The following sub-sections describe more clearly the GenioCSH methodology.

<sup>1</sup> [www.quidgest.com](http://www.quidgest.com)

### 3.1 The Team

During the execution of the block tests there should be, at least, two elements composing the development team present at the test. One (and only one, always) should be responsible for the execution of the individual tests, while the other(s) should be responsible for the execution of the proposed changes.

The elements responsible for the execution of the changes should be endowed with the capabilities for completing the tasks fast and correctly.

### 3.2 The Users

The ideal number of users present at a block test stays between 6 and 8 users. If the number of user is too small, under 4 users, there is the risk that some important changes are not detected by the users. In the opposite side, if the number of users is too big, over 12 users, the volume of changes proposed could be too massive, allowing two negative factor to happen: the elements responsible for the execution of the changes are overloaded with proposals to execute and, at the final individual tests, the amount of changes done to the initial system are so vast that the solution presented on those tests can already differ from the initially desired by the client and represent an unwanted solution, obviously with no interest to be tested.

The selection of the users present at the test should be done partially random, there with a higher emphasis on the users experienced and familiar with the feature tested. Doing tests with both kinds of users, familiar and not familiar, are in the interest of both parts since both represent different perspectives and ways of “facing” the related problem and the feature tested. Familiar users bring knowledge and experience into the execution of the task, while not familiar users bring to the test, the ability of abstraction to the common (usual) solutions to a problem, do not have “bad habits” and more easily produce new interpretations to the problem and to the solution. These not familiar users are, commonly the ones more capable of finding a new path from point A to point C, not passing through point B (as usually done).

### 3.3 The Tests

The tests can and should be done in all phases of the development, since the moment that there’s a functional version of the solution with a feature complete. With the use of RAD tools this baseline happens very early. Once again, the solution does not need to be completely functional, just the features covered on the test should be complete and with help options available (help menus, tooltips, etc).

The tests of a block test should be done continuously and contiguously, one test per user.

### 3.4 The Changes

Each user can propose as much changes as he consider necessary. From that list of changes, there should be iden-

tified and marked all changes already proposed by other users of the same block test that are already in execution. These shouldn’t be considered for late analysis, nevertheless, the user shouldn’t be warned of that fact.

After presenting all changes, the user should classify all changes regarding their importance level, on a scale from 1 (very little important) to 5 (very important). If the total number of changes is under 5, there shouldn’t be changes classified with the same importance. If the total number is over 5 changes, there should be repeated classifications, ensuring the fact that all levels of classification are used.

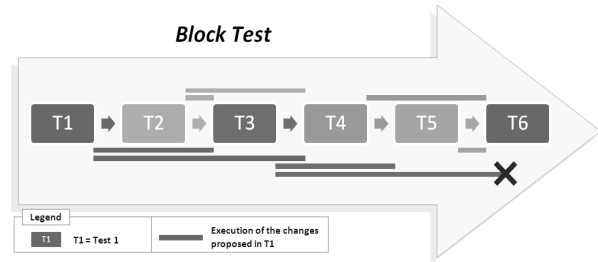


Figure 4 - The Changes

### 3.5 The Metrics

After the user classification of the proposed changes, the development team present at the block test should calculate the effective value of the proposed change. All changes already in development (proposed by a previous user) shouldn’t be considered.

To determine the change weight ( $CW$ ), three aspects are considered: the classification given by the user regarding the importance of the change ( $Im$ ), the familiarization of the user towards the feature/task tested ( $Fa$ ) and the number of times this change proposal was presented by the users ( $x$ ).

$$CW = \sum_{n=1}^x (Fa_n \cdot Im_n)$$

Figure 5 - The Metrics Formula

The value of the user familiarization towards the feature/task tested should be given by the development team, on a scale from 1 (not familiarized) to 3 (very much familiarized).

If a change proposal is presented several times, even being classified as with low importance, it gains more value each time it’s presented, could in this way, reach a value where it could be taken into execution. The more users feel the need for a change, the more it has value.

### 3.6 The Execution of the Changes

To execute the changes it is necessary human resources, and unfortunately, these resources, more concretely the development team present at the block test, are limited. The test duration is also short, representing a consequent short time available to execute the changes.

It is estimated that the number of simultaneous changes in execution is at most, the double of the human resources available in a certain moment (*max value*).

It's desired that the number of executed changes reach the maximum value possible, but what's important to prevent is the overload of the development team. It's desirable that the conclusion of a change is not compromised due to the overload. Since the GenioCSH intent is to complete as much changes as possible, keeping possible the addition of these changes into the next users testing solution is, without any doubt, preferable the conclusion of few changes instead of trying to conclude several changes and in the end, finishing none of them. It's better to have few changes complete and tested in next user tests, than to have none.

After each test, the change proposals should be categorized, calculated the change weight and checked if it should be (or not) put some changes into execution. This verification should take in consideration the number of changes actually in execution and should cover the change proposals from the last two user tests.

If the actual number of changes in execution is smaller than the *max value*, the changes with higher weight should start to be executed. Eventually if when reaching the *max value* of executions in simultaneous there are other changes (not in execution) with the same weight, these should also be executed.

If the actual number of changes in execution is higher than the *max value*, the only changes that should be started are the ones with weight higher than the maximum weight of the changes currently in execution.

### 3.7 The Validation

The validation of the changes is of high importance and it represents a key factor of the GenioCSH methodology.

After each block test the development team should present to the key users, project managers and supervisors, all the changes proposed by the users on their individual tests, where these should be analysed and validated.

The validation (or rejection) given to the changes will guarantee that the solution developed won't divert from the initial route. The changes proposed can represent an added value to the solution, but they can also represent a threat to the clients intention and objectives.

Only after the official approval, the changes can be considered as part of the solution, only then they will be effectively included in the solution.

## 4. IMPLEMENTATION

The methodology presented was implemented on the development process of two IS, both in use since the beginning of 2009.

The two projects (P1 and P2) are from different scopes and sizes. Regarding the appliance of GenioCSH in their implementation, in P1, GenioCSH was implemented since the beginning of the development, but while regarding P2, GenioCSH arose after the development started, being implemented a little less than half way thru his total development. Regarding their sizes, P1 consisted on a considerably smaller project than P2, as it can be verified on the following table.

Project	Objects	Code Lines
P1	27	119.315
P2	281	1.989.454

Table 1 - P1 and P2 objects and code lines

When GenioCSH started to be applied on the development of P2, it was suspected that the aspects related with the user interface and usability where underestimated creating lacks on the current solution.

## 5. RESULTS

As result of the application of GenioCSH there was registered a large quantity of change proposals.

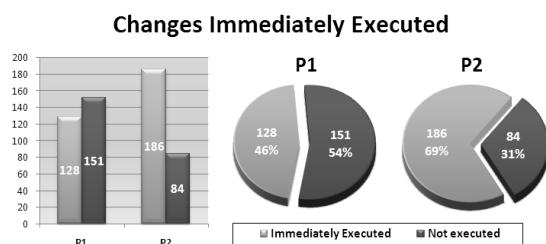
	Total
P1 - Usability	93
P1 - Functionality	186
P1 - Total	279
P2 - Usability	118
P2 - Functionality	155
P2 - Total	273

Figure 6 - Change proposals in P1 and P2

Nevertheless, these numbers don't represent that there was 279 (on P1) or 273 (on P2) errors on the IS developed since these number include several repeated proposal, proposals related with the hardware or network and proposals about inconsistencies detected with the objectives of the solution.

Even with the majority of change proposals been given in the P1, the project with more usability proposals was the P2. This higher volume of usability change proposals given in the P2 confirm the suspicious that the usability and user interface aspects were underestimated on the earlier developments on the P2.

As mentioned on the section 3, only some of the change proposals are considered to be executed during the block test.

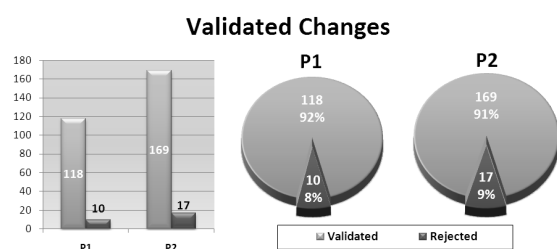


**Figure 7 - Changes immediately executed**

In both projects, a large quantity of changes were executed during the block test, but in the case of the P2, the percentage of executed changes is higher than in the P1. This happened due to the fact that the P2 had a lack of usability and user interface, and since GenioCSH wasn't applied from the beginning of the development, the solution had several "small" details that needed to be changed, without representing the need for big developments or resource consumptions. This way, it was possible to execute a higher percentage of the changes proposed in the P2.

At the end of each block test, the change were validated by the key users, project managers and supervisors. From the changes proposed, in both projects, more than 90% of the changes were validated.

In the case of the P1 only 10 changes were rejected, in the other hand, 17 changes were rejected in the P2.



**Figure 8 - Validated changes**

Apart from the results obtained during the development, some results were obtained after the IS got into production.

		Predicted	Registered	Difference
Number of development cycles	P1	10	9	-10%
	P2	14	13	-7,20%
Number of incidents and assistance requests occurred in the 1 <sup>st</sup> month	P1	40	21	-47,50%
	P2	50	36	-28%

**Table 2 - Predicted and Registered values**

Considering the predictions made by the development team, the results obtained were inferior in all cases.

The predicted number of development cycles decreased one cycle in both projects, but when comparing the registered number of incidents and assistance requests occurred in the 1<sup>st</sup> month, the reduction is drastic.

In the case of the P1, a project where GenioCSH was applied from the beginning of the development, the number of incidents and assistance requests occurred in the 1<sup>st</sup> month was less 47,5% than the predicted. This represents a reduction in almost half of the predicted.

## 6. CONCLUSIONS

The use of GenioCSH in the fast development of two projects proved to be possible to centre the development in the Human Being without increasing the development time. In the case of these two projects, the development time was considerably reduced.

It was also proved that the usability and user interface are in many cases underestimated. The application of GenioCSH represented an increase of the usability and functionality on each development cycle, representing a valid tool on the hand of the development teams.

The development of specific solutions requires from the development team a higher effort to know better the client, the users and their real needs. The introduction of GenioCSH in the development process, allows this effort to be soften since, through the several block tests, the solution is being refined towards the desired specific solution.

The action of the block tests together with the use of a RAD tool allows the immediate resolution of a vast volume of changes, also allowing them to be immediately tested and validated. In both projects, from the considerable volume of changes that were proposed and executed immediately, the acceptance level stayed above 90%, bringing us to the conclusion that the existence of individual tests with change proposals and immediate execution, in more than 90% of the cases, represents a positive and valuable issue to the IS. More than 90% of the times the users proposed a change, their change proposal was considered to be correct and was included in the final solution.

The application of GenioCSH in the development of two IS, produced positive effects not only in the development process but also after the development ended and the project started to be in production. It was registered a considerable smaller quantity of incidents or assistance requests than initially predicted, representing a smaller need for the development team to continue dispending time with these projects.

The use of the methodology GenioCSH here represents a positive aspect to both parts, development team and customer, since it helps and enrich the development process, resulting on the specific solution the customer desired.

As a final conclusion, the GenioCSH methodology proved to be a valid tool on the fast development centred in the Human Being, where development can be fast and where solutions are designed due to the specific need and objectives of the customer and their users.

### 6.1 Future Work

As future work it will be interest to do an analysis about the impact of, when executing the changes, being taken into consideration not only the changes proposed in the last two tests, but in all test (of that block test) done until that moment that haven't been taken into execution.

## 7. REFERENCES

- Nielsen, J. (1994).  
Usability Engineering  
Morgan Kaufmann - ISBN: 0-12-518406-9
- Nielsen, J., Mack, R. (1994)  
Usability Inspection Methods  
John Wiley and Sons - ISBN: 0-471-01877-5
- Medlock, M., Wixon, D., Terrano, M., Romero, R., Fultron, B. (2002)  
Using RITE method to improve products: a definition and a case study. Humanizing Design, Usability Professional's Association 2002 Annual Conference, <http://www.microsoft.com/usability/Playtest/Publications/Using%20the%20RITE%20Method%20to%20improve%20products.doc;%20a%20definition%20and%20a%20case%20study.doc>
- Dumas, J., Redish, J. (1999)  
A Pratical Guide to Usability Testing  
Intellect Books - ISBN: 1841500208
- Seffah, A., Gulliksen, J., Desmarais, M. (2005)  
Human-Centered Software Engineering - Integrating Usability in the Software Development Lifecycle  
Springer-Verlag - ISBN:140204027X
- Lim, K., Long, J. (1994)  
The MUSE Method for Usability Engineering  
Cambridge University Press - ISBN:0521474949
- Lim, K. (1996)  
Structured task analysis: an instantiation of the MUSE method for usability engineering  
Interacting with Computers #8, pages 31-50
- Nunes, D. (2001)  
Object Modeling for User-Centered Development and User Interface Design: The Wisdom Approach  
[http://dme.uma.pt/people/faculty/nuno.nunes/?page\\_id=20](http://dme.uma.pt/people/faculty/nuno.nunes/?page_id=20)

Aveiro, 13, 14 e 15 de Outubro de 2010



# Desenvolvimento Centrado no Utilizador em Ambiente de *Participatory Design*: uma experiência na área da saúde

Leonor Teixeira  
DEGEI / GOVCOPP, Univ de Aveiro  
3810-193 Aveiro  
lteixeira@ua.pt

Vasco Saavedra  
DEGEI - Univ. de Aveiro  
3810-193 Aveiro  
vsaavedra@ua.pt

Carlos Ferreira  
DEGEI - Univ Aveiro / CIO  
3810-193 Aveiro  
carlosf@ua.pt

Beatriz Sousa Santos  
DETI / IEETA – Univ. Aveiro  
Aveiro  
bss@ua.pt

## Sumário

*Este artigo descreve uma experiência de desenvolvimento de um Sistema de Informação (SI) na área da saúde, que beneficiou das práticas de Projecto Centradas no Utilizador (PCU), num ambiente de ‘Participatory Design’ (PD). Um conjunto de técnicas oriundas das disciplinas de Interacção Humano-Computador (IHC) e da Engenharia de Usabilidade (EU), combinadas com as tradicionais da Engenharia de Software (ES), permitiram chegar a uma solução eficaz e, do ponto de vista do utilizador, usável. Com base nesta experiência foi possível concluir, em conformidade com a literatura, os bons resultados normalmente conseguidos com a aplicação de abordagens PCU, em ambiente de PD. No entanto, e apesar destes bons resultados, a experiência com este projecto leva a acreditar que se não existir um certo controlo do procedimento por parte do analista, o processo de alteração de requisitos pode tornar-se infundável, devido ao ambiente favorável para a mudança rápida de ideias e, consequentemente, propício à reformulação dos requisitos.*

## Palavras-chave

*Sistemas de Informação na Saúde, Processo de Desenvolvimento, Envolvimento do utilizador, Projecto Centrado no Utilizador, Participatory Design, IHC, Engenharia de Usabilidade, Engenharia de Software.*

## 1. INTRODUÇÃO

O processo de desenvolvimento de Sistemas de Informação (SIs), embora com grandes contribuições da Engenharia de Software (ES), começou recentemente a integrar outras áreas de conhecimento de forma a abranger também os aspectos sociais e humanos associados à componente da Interacção. Estes SIs, para além da componente funcional ou *back-end*, comportam a componente de interface (*front-end*) constituindo esta, muitas vezes, o ‘espelho’ do SI aos olhos do utilizador. Como tal, durante o processo de desenvolvimento deste tipo de SIs, é necessário considerar, não só as especificações funcionais e técnicas, mas também todos os aspectos relacionados com a interface de utilizador, bem como com o processo de interacção. Do ponto de vista do utilizador, a separação destas componentes nem sempre é possível, devido ao sistema ser normalmente utilizado e avaliado como um todo. No entanto, conceptualmente, estas componentes podem ser projectadas usando conceitos de disciplinas distintas. Enquanto a componente técnica e funcional (*back-end*), relacionada com toda a estrutura interna que é definida a partir da especificação do cliente, é geral-

mente abordada na disciplina de ES; a componente da interface (*front-end*), que aborda os aspectos relacionados com a interacção entre o utilizador e o SI, está associada às disciplinas de Interacção Humano-Computador (IHC) e/ou Engenharia de Usabilidade (EU). Na realidade estas disciplinas trabalham processos diferentes de um mesmo produto.

No sentido de tentar fazer convergir o trabalho daquelas disciplinas, para um mesmo objectivo, são já vários os contributos existentes que tentam acomodar os métodos das diferentes áreas num mesmo processo. Por exemplo, o trabalho de [Harmelen01] refere o desenvolvimento de um método (OO&HCI – *Object-Oriented and Human-Computer Interaction*) que integra técnicas de modelação Orientadas a Objectos (OO), com os conceitos de IHC, no desenvolvimento de sistemas interactivos. Outra proposta de destaque nesta área é o método WISDOM (*Whitewater Interactive System Development with Object Models*) [Nunes00] que tenta integrar os métodos e técnicas da usabilidade nos processos tradicionais OO de desenvolvimento de software interactivo, combinando o melhor das duas disciplinas. A proposta de [Mayhew99],

concretizada através da *Framework 'Usability Engineering Lifecycle'*, resultou também de uma tentativa de redesenhar o processo de desenvolvimento de software em torno dos conhecimentos, métodos e actividades da EU. Enumeram-se ainda um conjunto de outros trabalhos que tentam, através de estudos empíricos e/ou modelos teóricos, compreender e explicar a dificuldade da integração plena destas duas disciplinas [Jerome05], [Sutcliffe05], havendo outros que chegam a propor pistas, especificando que técnicas de usabilidade, quando, e como, deverão ser integradas nas actividades de desenvolvimento de software [Ferre05].

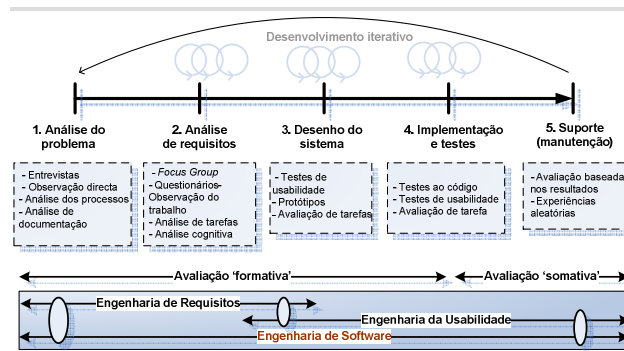
O desenvolvimento segundo os princípios de Projecto Centrados no Utilizador (PCU) surge também da tentativa de fundir as melhores práticas destas duas disciplinas (ES e IHC / EU), emergindo o utilizador como elemento central do desenvolvimento, a par das tarefas e do contexto de utilização [Maguire01]. Estas práticas podem envolver o utilizador de uma forma activa ou passiva. No caso do utilizador ser envolvido de forma activa, isto é, ser integrado na equipa de projecto e participar nas decisões, diz-se estar perante um ambiente de *Participatory Design* (PD).

A presente artigo descreve uma experiência de desenvolvimento de um SI na área da saúde, que beneficiou de práticas de PCU, num ambiente de PD. Para uma melhor explanação da experiência e das escolhas feitas em torno do processo de desenvolvimento, estruturou-se o trabalho em 4 secções. A presente secção foi dedicada a uma breve introdução. A secção 2 reporta, com base na literatura e com foco na área da saúde, algumas abordagens de desenvolvimento que emergiram do factor humano, destacando-se a este nível as práticas iterativas e incrementais, centradas no utilizador, e com processos de avaliação 'formativa'. A secção 3 descreve a experiência de desenvolvimento de um SI na saúde, cujo processo de desenvolvimento foi inspirado nas práticas de PCU em ambiente de PD. Finalmente uma breve conclusão na secção 4.

## 2. ABORDAGENS DE DESENVOLVIMENTO EMERGENTES DO FACTOR HUMANO

Nos últimos anos, particularmente na área da saúde, as metodologias de desenvolvimento de SIs têm dado grande ênfase à aplicação de práticas de desenvolvimento iterativas e incrementais, com processo de avaliação ao longo de todo ciclo [Brender06], [Kushniruk02] [Kushniruk04]. Por outro lado, importantes desenvolvimentos na área da ES têm emergido das variações aplicadas aos métodos tradicionais, tendo como base os princípios do projecto iterativo e centrado no utilizador [Samaras05], [Seffah05], [Seffah05a], [Zimmermann07]. De facto a literatura confirma que as metodologias de análise e projecto centradas no utilizador, com foco central na avaliação contínua, em processos iterativos de avaliação 'formativa', têm sido das mais procuradas no desenvolvimento de SIs interactivos de uma forma geral, e no sector da saúde em particular [Johnson05], [Rinkus05], [Samaras05], [Zhang05], [Zhang02]. Por

exemplo, a *Framework* de desenvolvimento apresentada por [Kushniruk02] e [Kushniruk04] retrata uma estrutura que relaciona as principais técnicas e métodos de avaliação com as respectivas etapas que compõem o ciclo de desenvolvimento do SI, combinando um conjunto de técnicas vindas da IHC/EU, com os métodos tradicionais da ES, fig.1.



**Fig. 1: Framework de desenvolvimento de SI que integra técnicas e métodos da IHC e da EU ao longo de um processo iterativo (adaptado [Kushniruk02] e [Kushniruk04]).**

A Engenharia de Requisitos (ER), embora considerada uma subárea da ES, emergiu da necessidade de incluir os aspectos humanos e sociais, nas fases preliminares do ciclo de desenvolvimento, realçando de certa forma a importância que o envolvimento dos utilizadores tem na definição de SIs usáveis e com qualidade [Zimmermann07].

### 2.1 Princípios de Projecto Centrados no Utilizador e Participatory Design

A comunidade de IHC tem desenvolvido uma variedade de abordagens e métodos de forma a envolver os utilizadores e a integrar os conceitos oriundos da EU [Nielsen93], [Mayhew99], [Preece02]. Neste âmbito surgiram também as metodologias de Projecto Centradas no Utilizador (PCU), complementares às abordagens tradicionais, acrescentando um conjunto de princípios que colocam o utilizador no centro das atenções [Norman86]. Para além do utilizador, bem como suas características e tarefas, o contexto de utilização emerge como requisito importante no âmbito destas metodologias [Hackos99], tendo como principal objectivo a criação de sistemas após conhecimento sólido das características dos utilizadores e tarefas por eles realizadas [Johnson05]. Segundo estes autores, o resultado de um bom projecto traduz-se na sua facilidade de utilização, no aumento da aceitação do sistema por parte do utilizador, no aumento da satisfação e da produtividade do utilizador, na diminuição de erros através da utilização do sistema e ainda, na diminuição da curva de aprendizagem em relação à utilização do novo sistema.

São já vários os trabalhos que seguem uma abordagem metodológica segundo os princípios do PCU [Ji-Ye05], [Rinkus05], [Samaras05], [Seffah05a], [Zimmermann07], tendo sido alguns destes desenvolvidos no âmbito de projecto de SIs no sector da saúde. Estas abordagens, não se

regem por métodos formais, mas antes, por um conjunto de técnicas e princípios que colocam o utilizador no centro do desenvolvimento [Norman86].

Técnicas oriundas das Ciências Sociais e Cognitivas, tais como o questionário, a entrevista, a análise documental, técnicas etnográficas ou de observação directa, são das mais utilizadas para a recolha de dados. De entre as várias referidas, alguns autores apontam a técnica da observação directa como aquela que mais se adequa à recolha do grande montante de dados, normalmente registado nas fases iniciais do processo de desenvolvimento [Johnson05]. É uma abordagem etnográfica que envolve a observação do utilizador no seu próprio ambiente de trabalho [Denzin05] e que se enquadra, segundo a perspectiva de [Nuseibeh00] dentro das técnicas contextuais. Com o uso destas, é possível extrair determinados detalhes que, de forma natural, enriquecem os dados e que seriam impossíveis de obter, por exemplo, através das técnicas tradicionais, como o questionário ou a entrevista.

Apesar das técnicas aplicadas segundo o princípio de PCU colocarem o utilizador no centro do desenvolvimento, tentando compreender as suas necessidades e limitações, na maioria das vezes não integram o utilizador de forma activa na equipa, nem o envolvem nas decisões do projecto. Neste caso está-se perante uma abordagem de PCU, numa filosofia de projecto “não participado”.

*Participatory Design (PD)* é considerado actualmente uma variante do PCU [Almaghairbe10], [Odueke10], [Incecik10], [Alazwari10] onde os utilizadores e/ou outros *stakeholders* são activamente integrados na equipa de desenvolvimento, de forma a serem envolvidos nas decisões associadas ao projecto. Esta abordagem teve origem na Escandinávia, com o nome de Projecto Cooperativo, enfatizando a importância do envolvimento dos utilizadores em decisões relacionadas com os sistemas computacionais que tivessem impacto no seu trabalho [Carroll96], [Kensing98], [Bodker96]. No entanto, e num ambiente de PD, o envolvimento do utilizador deverá ser mediado através de técnicas de fácil utilização e compreensão por parte dos utilizadores, destacando-se a este nível os *mock-ups*, os protótipos, as técnicas de *brainstorming*, e os jogos interactivos e de cenários.

### 3. CASO PRÁTICO: DESENVOLVIMENTO DE UM SISTEMA DE INFORMAÇÃO NA ÁREA DA SAÚDE COM BASE NUMA ABORDAGEM ‘PARTICIPATORY DESIGN’

No âmbito do desenvolvimento de SIs interactivos na área da saúde, a realidade encontrada é ainda um pouco distinta daquela que se verifica na indústria tradicional. A cultura que predomina neste domínio é a de treinar as pessoas para se adaptarem às tecnologias e não adaptar a tecnologia às características das pessoas [Zhang05]. No entanto, e devido aos problemas que daí advêm, começaram a aparecer trabalhos neste domínio com contribuições metodológicas que tentam ‘inverter’ aquele paradigma, dando também atenção à componente humana e social dos SIs, e não apenas aos aspectos técnicos e tec-

nológicos, [Coiera07], [Berg98], [Berg99]. Nesta linha de pensamento encontram-se ainda outros trabalhos que tentam colocar o utilizador como o elemento central no processo de desenvolvimento de SIs interactivos, argumentando ser um dos passos mais importantes para a obtenção de um produto com qualidade [Beuscart-Zéphir05], [De Rouck08], [Staccini01].

O presente artigo descreve a abordagem de desenvolvimento de um SI na área da saúde, mais especificamente um Sistema de Gestão de Informação Clínica de apoio à prática clínica da hemofilia (designado *hemo@care*), segundo os princípios de PCU, num ambiente de PD. A escolha desta abordagem baseou-se na análise das melhores práticas de desenvolvimento dos SIs na saúde reportadas pela literatura, tendo resultado de uma combinação de técnicas e métodos oriundos da IHC e EU, com as tradicionais metodologias de ES.

#### 3.1 Breve descrição do *hemo@care*

O *hemo@care*, é um SI com base na Web criado com o objectivo de apoiar a gestão da informação gerada no âmbito da prática clínica da hemofilia. Foi desenvolvido em estreita colaboração com o Serviço de Hematologia do Centro Hospitalar de Coimbra (SH\_CHC), Centro de Tratamento de Hemofilia (CTH) que dá assistência aos doentes hemofílicos da zona centro da País.

A hemofilia é uma doença crónica, de prevalência genética que se caracteriza pela ausência, ou uma forte carência, de um dos factores da coagulação no sangue. É uma patologia rara, pelo facto de afectar um pequeno número de pessoas, em comparação com a população em geral (1 caso em cada 10.000 pessoas). Actualmente não é conhecida a cura para a hemofilia, mas existem formas de tratamento que permitem minimizar os danos causados pela doença, possibilitando aos doentes uma qualidade de vida muito próximo do padrão normal. Na realidade, uma das particularidades da doença prende-se precisamente com estes tratamentos que acompanham a vida do paciente, e que podem ser feitos em regime domiciliário, vindo o seu registo a constituir protocolo obrigatório. Todos os dados relativos ao tratamento, episódio hemorrágico e fármacos utilizados têm que ser registados e, posteriormente, enviados para o CTH, para fins de controlo. O *hemo@care*, para além de apoiar a gestão dos dados resultantes da prática clínica da hemofilia, pretende também apoiar o processo de registo e de envio dos dados dos tratamentos domiciliários para os CTH onde o paciente está a ser acompanhado. Para tal, foi criado um perfil ‘Paciente’ com acesso a uma área do sistema através da Internet, que lhe permite registar todos os dados relacionados com os tratamentos efectuados em regime domiciliário.

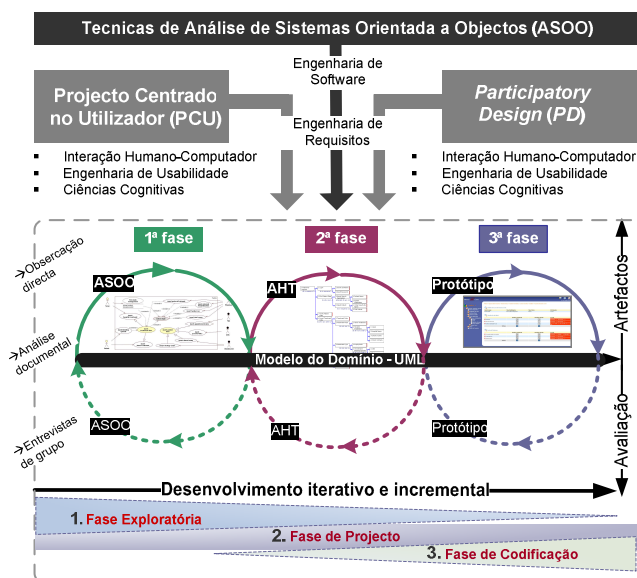
No caso dos tratamentos serem efectuados em regime hospitalar, o registo terá que ser efectuado pelo enfermeiro responsável pela administração do tratamento, estando esta funcionalidade também disponível no perfil ‘Enfermeiro’. Este tem ainda acesso a um conjunto de funcionalidades que lhe permitem efectuar e controlar a gestão de

*stocks* dos fármacos utilizados nos tratamentos, que até então eram feitos manualmente.

O ‘Médico’, também actor do sistema, a par do ‘Enfermeiro’ e do ‘Paciente’, tem acesso a uma área com um conjunto alargado de funcionalidades que lhe permitem registar, actualizar e pesquisar informação de natureza clínica. Uma descrição mais detalhada do *hemo@care* pode ser encontrada em [Teixeira08].

### 3.2 Descrição da abordagem adoptada no desenvolvimento do *hemo@care*

A abordagem que esteve na origem do desenvolvimento do *hemo@care* foi inspirada em técnicas de PCU num ambiente de *Participatory Design* (PD), em torno de um processo iterativo e incremental, suportado pelos métodos tradicionais de Análise de Sistemas Orientado a Objectos (ASOO), fig.2.



**Fig.2: Abordagem de desenvolvimento iterativa e incremental com base nos princípios de PCU em ambiente de PD.**

Com esta abordagem foi possível confrontar os potenciais utilizadores com artefactos resultantes da evolução iterativa do projecto, e com base nisso promover a discussão e obter o *feedback*, de forma a validar os requisitos previamente encontrados e, conseqüentemente identificar novos requisitos.

As técnicas de PCU contribuíram com mecanismos de recolha e avaliação de dados apropriados à compreensão das necessidades dos utilizadores, no contexto específico de trabalho em que estavam inseridos.

Com o ambiente de PD utilizaram-se mecanismos que permitiram inserir um grupo de potenciais utilizadores na equipa de desenvolvimento, contribuindo estes activamente para a (re)formulação dos requisitos do sistema. Dois médicos, dois enfermeiros e um paciente fizeram parte da equipa de projecto, sendo a sua participação efectuada em momentos oportunos do processo, no âmbito de reuniões programadas.

Por último, a componente de ASOO, suportada através de linguagem UML, foi utilizada para documentar os resultados que surgiam ao longo do processo de desenvolvimento iterativo e, sempre que necessário, comunicá-los através de uma linguagem formal, aos elementos da equipa responsáveis pela codificação (programadores).

A Engenharia de Requisitos, embora considerada uma subárea da ES, recorre às Ciências Sociais e Cognitivas para fundamentar a teoria usada aproximando-se, de alguma forma, das práticas tidas nas abordagens de PCU e PD [Sommerville07].

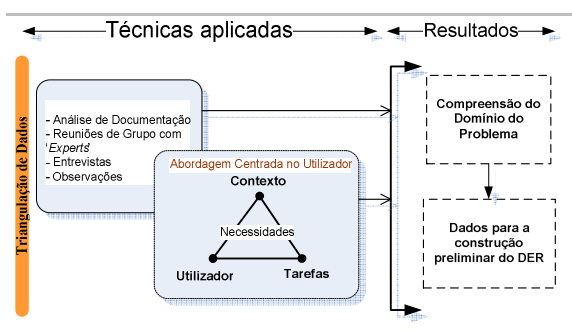
As técnicas e métodos utilizados ao longo do processo de desenvolvimento do *hemo@care* foram agrupados em três grandes fases (fase exploratória, fase de projecto e fase de codificação) tendo, cada uma delas, um conjunto de *inputs* e *outputs*, bem com contribuições de diferentes disciplinas. A fronteira entre as fases não é explícita, havendo sobreposição de actividades que advêm da abordagem iterativa adoptada ao longo do ciclo de desenvolvimento, fig.2. Seguidamente será feita uma breve descrição de cada fase, bem como as principais técnicas e métodos que contribuíram para os resultados de cada uma delas segundo os princípios de PCU em ambiente de PD.

#### 3.2.1 Fase Exploratória

Esta fase é aquela que caracteriza o trabalho nas primeiras etapas de desenvolvimento, onde se tentam obter os primeiros dados e se tem conhecimento do domínio do problema. Alguns autores afirmam que esta etapa é de extrema complexidade sendo nalgumas situações levada a cabo por profissionais de Marketing, quando deveria ser desenvolvida por especialistas em Engenharia de Requisitos ou Ergonomistas [Samaras05]. No sector da saúde, e pelo facto dos processos serem complexos com profissionais altamente especializados, as actividades relacionadas com esta fase ocupam uma posição crítica, tendo um impacto directo na qualidade do resultado final. Enquanto uma correcta análise do problema e identificação dos requisitos podem constituir um factor determinante para a obtenção de bons resultados, uma deficiência na obtenção dos mesmos poderá conduzir a resultados fracassados. Segundo a visão de [Samaras05] este tipo de falhas não é apenas atribuída à insuficiência na identificação dos potenciais utilizadores e suas necessidades, mas também ao facto de se partir do pressuposto de que já se conhecem bem as necessidades daqueles, admitindo assim a conclusão prematura do trabalho desta fase. Na realidade, e principalmente no domínio da saúde, os requisitos dos utilizadores são bastante voláteis e extremamente mutáveis, dependendo de vários factores, nomeadamente do contexto em que são adquiridos. Em comparação com outros sectores, neste (especialmente na prestação dos cuidados de saúde) existem vários aspectos específicos que tornam o trabalho desta fase particularmente difícil [Cysneiros02], [Garde06]. Desses aspectos, realçam-se os que interferem com o processo de comunicação entre os membros da equipa e os especialistas do domínio, como por exemplo:

- O vocabulário utilizado ao nível do domínio da saúde que geralmente não é do conhecimento do engenheiro de requisitos, requerendo um maior esforço por parte deste;
- O conhecimento dos especialistas do domínio do problema que, para além de bastante especializado e complexo, está associado a estruturas tácitas, o que dificulta a sua transmissão e/ou comunicação através das técnicas tradicionais de obtenção de requisitos.

Nesta fase foram utilizadas técnicas que, segundo a terminologia de [Nuseibeh00], se podem classificar no âmbito das técnicas tradicionais (ex.: análise de documentação), das técnicas de grupo (ex.: reuniões/ entrevista de grupo) e das técnicas contextuais (ex.: observação directa). Estas técnicas foram aplicadas segundo os princípios de PCU em torno dos conceitos de tarefa, utilizador e contexto, fig. 3. A sua aplicação foi combinada no sentido de tirar partido da triangulação de dados [Cysneiros02], mais concretamente, completar a recolha com novos dados obtidos através de novas abordagens (técnicas), ao mesmo tempo que se validavam os previamente recolhidos.



**Fig.3: Técnicas aplicadas e resultados obtidos na fase exploratória.**

Os *outputs* desta fase resumem-se na compreensão do domínio do problema, e num conjunto de dados que permitiram construir a versão preliminar do Documento de Especificação de Requisitos (DER).

Após esta fase, que decorreu nos cerca de primeiros 3 meses do ciclo de desenvolvimento, passou-se à seguinte a que, pela natureza do trabalho predominante, se deu o nome de fase de projecto.

### 3.2.2 Fase de Projecto

A fase de projecto acrescentou à primeira uma notação de suporte que, através da sua sintaxe e semântica, deram forma ao modelo conceptual. Através de modelos é possível representar e visualizar a abstracção completa de uma realidade complexa, para além de ser relativamente fácil comunicar com os utilizadores e *stakeholders* e, consequentemente, validar os resultados previamente encontrados.

Três técnicas diferentes (ASOO, AHT e Prototipagem) de representação e avaliação foram utilizadas, sendo os artefactos construídos num ambiente de *Participatory Design* (PD).

- **Análise de Sistemas Orientado a Objectos (ASOO)** – a primeira técnica aplicada, foi escolhida no âmbito das disponíveis pela disciplina de ES, baseou-se na tradicional análise de sistemas com base num paradigma de Orientação a Objectos (OO), sendo a UML a notação adoptada. Com base nesta técnica foram construídos os modelos, destacando-se os diagramas de *Use Cases* e de Classes. O diagrama de Classes foi utilizado para documentar os resultados em termos de modelo de dados, enquanto que o diagrama de *Use Cases*, para além de documentar os actores e interacções destes com o sistema, foi extensivamente utilizado para validar as funcionalidades com os utilizadores, bem como para encontrar funcionalidades em falta (novas). No âmbito de reuniões de grupo, os potenciais utilizadores envolvidos no PD eram confrontados com os artefactos no sentido de os discutirem, criticarem e apresentarem sugestões de reformulação. Dadas as restrições da notação (da linguagem UML), pouco perceptível por pessoas sem qualquer *background* tecnológico, na apresentação dos modelos utilizou-se uma linguagem adequada, tendo convertido os *use-cases* em breves relatos ou narrativas (*user stories*) que tentavam descrever o que o sistema fazia. O *feedback* dos utilizadores era também baseado nos relatos e, posteriormente, convertido para *use-cases*. Após 3 iterações, não se verificando mais evoluções no modelo, passou-se para a técnica seguinte.

- **Análise de Tarefas** – a segunda técnica escolhida, muito utilizada na área da IHC, baseou-se na Análise de Tarefas, mais especificamente a Análise Hierárquica de Tarefas (AHT) [Annett04]. Esta foi construída com base no resultado da fase anterior, tendo cada *use-case* gerado um modelo AHT. Nesta fase nem todos os *use-cases* foram convertidos em modelos, mas apenas os mais complexos e que envolviam um maior número de operações. O objectivo desta fase era validar as funcionalidades previamente encontradas, ao mesmo tempo que se tentava perceber o modelo mental do utilizador de forma a encontrar a sequência de apresentação da informação no âmbito de uma funcionalidade. Esta técnica foi também iterada três vezes, tendo-se passado à aplicação da técnica seguinte, após estagnação do modelo.

- **Prototipagem** – foi o terceiro método aplicado, muito conhecido no âmbito da IHC, mas também bastante utilizado no âmbito da Engenharia de Requisitos (ER) [Pressman05]. Tendo em conta o conceito que se pretendia testar, optou-se por um protótipo de baixa fidelidade, desenvolvido segundo o método *'throw-away'*, mais concretamente um protótipo horizontal [Andriole94]. Com base nas funcionalidades encontradas e na sequência de operações percebida com a técnica anterior, foram criados *mock-ups*, com recurso a HTML. Estes *mock-ups*, colocados numa sequência lógica de acordo com a sequência identificada com a AHT, pretendiam simular a interface de utilizador da aplicação de forma a serem validados pelo utilizador. A



reformulação das interfaces foi feita na presença dos utilizadores, num ambiente de *PD*, tendo o modelo estabilizado após duas iterações.

As três técnicas foram aplicadas de forma iterativa e incremental, pela ordem em que foram apresentadas. A tabela 1 apresenta algumas vantagens que se destacaram na fase de projecto com a aplicação de cada uma das técnicas.

Técnica	Vantagem verificadas com a aplicação da técnica
ASOO (UML)	<ul style="list-style-type: none"> <li>Representação OO adequada à natureza evolutiva do projecto.</li> </ul>
Análise de Tarefas (AHT)	<ul style="list-style-type: none"> <li>Técnica adequada à compreensão do modelo mental dos utilizadores;</li> <li>Técnica adequada à identificação de sub-funcionalidades e ordem da execução dentro das funcionalidades de mais alto nível;</li> <li>Notação de fácil compreensão e adequada a um processo de comunicação entre pessoas com <i>background</i> de conhecimento diferentes.</li> </ul>
Prototipagem (P. Horizontal)	<ul style="list-style-type: none"> <li>Técnica adequada à captação da linguagem para a representação das funcionalidades;</li> <li>Técnica adequada à captação de requisitos emergentes;</li> <li>Modelo de fácil compreensão proporcionando, como tal, maior entusiasmo de participação por parte dos utilizadores.</li> </ul>

**Tabela 1: Algumas vantagens verificadas com a aplicação das técnicas.**

### 3.2.3 Fase de Codificação

A fase de codificação, em que o modelo conceptual resultante da fase anterior foi convertido em código executável, baseou-se numa abordagem de desenvolvimento ágil, mais concretamente o *eXtreme Programming* (XP) [Beck99]. Esta abordagem foi escolhida pelo facto do desenvolvimento ser feito através de iterações, obtendo *feedback* rápido por parte do utilizador, através de testes às versões do software desenvolvido. De certa forma é uma abordagem bastante consistente com as práticas de PCU, apresentando características de protótipo evolutivo e, por isso, adequada ao desenvolvimento de projectos que necessitem de lidar com mudanças rápidas e imprevisíveis dos requisitos, como acontece nos SIs da área da saúde.

### 3.3 Avaliação ‘somativa’ do *hemo@care*: teste de observação de tarefas, avaliação heurística e testes em ambiente real

Após terminado o desenvolvimento, segundo uma abordagem centrada no utilizador em ambiente de *PD*, o *hemo@care* foi submetido, numa primeira fase, a uma avaliação de usabilidade (avaliação ‘somativa’) através de testes de observação de tarefas e avaliação heurística. Com esta avaliação pretendia-se obter *feedback* de um pequeno grupo de utilizadores com conhecimentos em IHC, relativamente a algumas características da interface, bem como compreender alguns atributos ao nível da facilidade de aprendizagem e utilização da interface da aplicação. No que respeita ao teste de observação de tarefas, a interface do paciente foi privilegiada, não só pelo facto de ser aquela que teve menor representatividade ao longo

de todo o processo de desenvolvimento em termos de utilizadores participantes no processo (apenas um paciente fez parte da equipa numa lógica de *PD*), como também por se tratar de um perfil que pode contemplar utilizadores com as mais diversas características e níveis de conhecimento. Sendo assim, uma lista de 15 tarefas do perfil ‘Paciente’ foram escolhidas para avaliação, algumas delas comuns aos perfis ‘Enfermeiro’ e ‘Médico’.

Este tipo de avaliação implica a colaboração de utilizadores e observadores. Neste caso particular, cada observador, previamente treinado, acompanhava um utilizador e, com base num guião, registava todos os dados relativos ao comportamento e desempenho do utilizador, enquanto este realizava as tarefas que faziam parte de uma lista pré-definida.

Na mesma altura foi também realizada uma avaliação heurística por dois avaliadores, com base nas dez heurísticas de usabilidade propostas por J. Nielsen [Nielsen05].

Os resultados dos testes de observação de tarefas e da avaliação heurística revelaram-se surpreendentes e foram bastante positivos. Pequenas melhorias foram implementadas na interface do ‘Paciente’, quer relativamente ao aspecto gráfico, quer ao nível de algumas funcionalidades que, pela sua importância, apareciam num lugar de pouco destaque. No entanto, aponta-se como limitação ao teste de observação de tarefas, o facto do grupo de observados poderem não ser representativos dos potenciais utilizadores, uma vez que tinham conhecimentos de IHC, mas desconheciam o domínio do problema.

Numa segunda fase, fizeram-se testes em ambiente real com utilizadores reais. Estes testes contemplaram maioritariamente os perfis ‘Médico’ e ‘Enfermeiro’, tendo consistido em avaliações livres com dados reais e em ambiente de trabalho normal. Os utilizadores usavam a aplicação de teste, simulando a utilização de uma solução final, ao mesmo tempo que registavam todas as dificuldades que sentiam com a entrada de dados e/ou visualização da informação. Este tipo de testes, ainda que não utilize um protocolo formal, requer o envolvimento pleno dos utilizadores, em muitos casos, sem a presença de qualquer membro da equipa de desenvolvimento. O registo de todas as dificuldades sentidas por parte do utilizador é bastante relevante, para se perceberem os reais problemas de usabilidade e, posteriormente, se incrementarem as melhorias correspondentes. Neste caso particular, este teste permitiu implementar pequenas melhorias no *hemo@care*, encontrando-se actualmente ainda instalado no SH\_CHC para esse fim (fase de testes). Os resultados verificados até então, ainda que contribuíssem para o incremento de pequenas melhorias, permitiram também confirmar a escolha acertada ao nível da abordagem adoptada durante todo processo de desenvolvimento do *hemo@care*.

## 4. CONCLUSÃO

O presente artigo retratou o processo de desenvolvimento de um SI na área da saúde, que beneficiou das práticas de PCU, em ambiente de *PD*. Esta abordagem, em confor-

midade com a literatura, revelou-se como uma das mais adequadas ao desenvolvimento de SI interactivos, inseridos em ambientes de incerteza no que respeita ao levantamento de requisitos, como aquele que caracteriza o domínio da saúde.

Na realidade, desenvolver SIs interactivos eficientes, fiáveis, fáceis de utilizar e com curvas de aprendizagem baixas é actualmente um desafio que a indústria de desenvolvimento de software enfrenta. No sector da saúde, área caracterizada pela existência de processos extremamente mutáveis e complexos e onde o sucesso dos seus SIs depende fortemente da reflexão tida ao nível dos aspectos sociais e humanos, as abordagens tradicionais de desenvolvimento terão naturalmente melhores resultados, quando integradas com conhecimentos vindos de disciplinas que envolvam a compreensão daqueles factores, tais como da IHC e da EU.

Os factores humanos e sociais no âmbito do desenvolvimento de sistemas envolvem a descoberta de informação respeitante ao comportamento humano, que pode ser mais facilmente conseguido com abordagens de desenvolvimento segundo aos princípios de PCU, com especial destaque para ambientes de *Participatory Design (PD)*. A experiência aqui descrita confirma este facto, tendo-se obtido bons resultados pelo facto de se envolverem os utilizadores durante todo o processo de desenvolvimento. No entanto, e apesar destes bons resultados, a experiência com este projecto leva a acreditar que se não existir um certo controlo do procedimento por parte do analista, o processo de alteração de requisitos pode tornar-se infundável, devido ao ambiente favorável para mudanças rápidas de ideias e, consequentemente, propício à reformulação dos requisitos.

## 5. AGRADECIMENTOS

Este trabalho foi conseguido, graças à cooperação dos profissionais dos Serviços de Hematologia do Centro Hospitalar de Coimbra, com um contributo especial da Doutora Natália Martins e do Doutor Ramon Salvado; aos pacientes daquele Centro, e à preciosa colaboração da Associação Portuguesa dos Hemofílicos (APH), para os quais vão os nossos agradecimentos.

## 6. REFERÊNCIAS

- [Alazwari10] Alazwari, S. (2010). "Participatory Design: Methodologies and Techniques." *Proceedings of Designing Usable Systems*: 35-38.
- [Almaghairbe10] Almaghairbe, R. (2010). "Overview in Participatory Design." *Proceedings of Designing Usable Systems*: 31-34.
- [Andriole94] Andriole, S. J. (1994). "Fast, cheap requirements prototype, or else!" *Software, IEEE* 11(2): 85-87.
- [Annett04] Annett, J. (2004). *Hierarchical Task Analysis. The Handbook of Task Analysis for Human-Computer Interaction*. D. Diaper and N. A. Stanton. Mahwah, New Jersey, Lawrence Erlbaum Associates: 67-82.
- [Beck99] Beck, K. (1999). "Embracing Change with Extreme Programming." *Computer* 32(10): 70-77.
- [Berg99] Berg, M. (1999). "Patient care information systems and health care work: a sociotechnical approach." *International Journal of Medical Informatics* 55(2): 87-101.
- [Berg98] Berg, M., C. Langenberg, et al. (1998). "Considerations for sociotechnical design: experiences with an electronic patient record in a clinical context." *International Journal of Medical Informatics* 52(1-3): 243-251.
- [Beuscart-Zéphir05] Beuscart-Zéphir, M.-C., F. Anceaux, et al. (2005). "User-centred, multidimensional assessment method of Clinical Information Systems: a case-study in anaesthesiology." *International Journal of Medical Informatics* 74(2-4): 179-189.
- [Bodker96] Bodker, S. (1996). "Creating conditions for participation: conflicts and resources in systems development." *Hum.-Comput. Interact.* 11(3): 215-236.
- [Brender06] Brender, J. (2006). *Handbook of Evaluation Methods for Health Informatics*, Elsevier Inc.
- [Carroll96] Carroll, J. M. (1996). "Encountering Others: Reciprocal Openings in Participatory Design and User-Centered Design." *Human-Computer Interaction* 11(3): 285 - 290.
- [Coiera07] Coiera, E. (2007). "Putting the technical back into socio-technical systems research." *International Journal of Medical Informatics* 76(Supplement 1): S98-S103.
- [Cysneiros02] Cysneiros, L. (2002). *Requirements Engineering in the Health Care Domain*. 10th Anniversary IEEE Joint International Conference on Requirements Engineering. Washington - USA, IEEE Computer Society: 350-356.
- [De Rouck08] De Rouck, S., A. Jacobs, et al. (2008). "A methodology for shifting the focus of e-health support design onto user needs: A case in the homecare field." *International Journal of Medical Informatics* 77: 589-601.
- [Denzin05] Denzin, N. K. and Y. S. Lincoln (2005). *The SAGE handbook of qualitative research*. Thousand Oaks (CA), Sage Publications.
- [Ferre05] Ferre, X., N. Juristo, et al. (2005). Which, When and How Usability Techniques and Activities Should Be Integrated. *Human-Centered Software Engineering - Integrating Usability in the Software Development Lifecycle*. A. Seffah, J. Gulliksen and M. C. Desmarais. Netherlands, Springer. 8: 173-200.
- [Garde06] Garde, S. and P. Knaup (2006). "Requirements engineering in health care: the example of chemotherapy planning in paediatric oncology." *Requirements Engineering* 11(4): 265-278.
- [Hackos99] Hackos, J. T. (1999). *User and task analysis for interface design*. New York, John Wiley.
- [Harmelen01] Harmelen, M. v. (2001). *Interactive system design using OO&HCI methods*. Object modeling and user interface design: designing interactive systems, Addison-Wesley Longman Publishing Co., Inc.: 365-427.

- [Incecik10] Incecik, R. E. (2010). "Participatory Design Assessment Regarding to Benefits & Limitations." *Proceedings of Designing Usable Systems*: 39-42.
- [Jerome05] Jerome, B. and R. Kazman (2005). *Surveying the Solitudes: an investigation into the relationship between human computer interaction and software engineering in practice*. Human-Centered Software Engineering - Integrating Usability in the Software Development Lifecycle. A. Seffah, J. Gulliksen and M. C. Desmarais. Netherlands, Springer. 8: 59-70.
- [Ji-Ye05] Ji-Ye, M., V. Karel, et al. (2005). "The state of user-centered design practice." *Communications of the ACM* 48(3): 105-109.
- [Johnson05] Johnson, C. M., T. R. Johnson, et al. (2005). "A user-centered framework for redesigning health care interfaces." *Journal of Biomedical Informatics* 38(1): 75-87.
- [Kensing98] Kensing, F. and J. Blomberg (1998). "Participatory Design: Issues and Concerns." *Computer Supported Cooperative Work (CSCW)* 7(3): 167-185.
- [Kushniruk02] Kushniruk, A. (2002). "Evaluation in the design of health information systems: application of approaches emerging from usability engineering." *Computers in Biology and Medicine* 32(3): 141-149.
- [Kushniruk04] Kushniruk, A. and V. Patel (2004). "Cognitive and usability engineering methods for the evaluation of clinical information systems." *Journal of Biomedical Informatics* 37(1): 56-76.
- [Maguire01] Maguire, M. (2001). "Methods to support human-centred design." *International Journal of Human-Computer Studies* 55(4): 587-634.
- [Mayhew99] Mayhew, D. J. (1999). *The usability engineering lifecycle*. San Francisco (CA), Morgan Kaufman.
- [Nielsen93] Nielsen, J. (1993). *Usability engineering*. Boston, AP Professional.
- [Nielsen05] Nielsen, J. (2005). "Ten Usability Heuristics." [http://www.useit.com/papers/heuristic/heuristic\\_list.html](http://www.useit.com/papers/heuristic/heuristic_list.html).
- [Norman86] Norman, D. A. and S. W. Draper (1986). *User Centered System Design; New Perspectives on Human-Computer Interaction*, L. Erlbaum Associates Inc.
- [Nunes00] Nunes, N. J. and J. F. Cunha (2000). "Wisdom: A Software Engineering Method for Small Software Development Companies." *IEEE Software* 17(5): 113-119.
- [Nuseibeh00] Nuseibeh, B. and S. Easterbrook (2000). *Requirements engineering: a roadmap*. *Proceedings of the IEEE International Conference on Software Engineering*. Limerick, Ireland, ICM Press: 35-46.
- [Odueke10] Odueke, A. (2010). "Participatory Design and Co-design: Challenges and Considerations." *Proceedings of Designing Usable Systems*: 43-46.
- [Preece02] Preece, J., Y. Rogers, et al. (2002). *Interaction Design: Beyond Human-Computer Interaction*. New York (NY), John Wiley.
- [Pressman05] Pressman, R. S. (2005). *Software Engineering: a practitioner's approach*. Boston (MA), McGraw-Hill.
- [Rinkus05] Rinkus, S., M. Walji, et al. (2005). "Human-centered design of a distributed knowledge management system." *Journal of Biomedical Informatics* 38(1): 4-17.
- [Samaras05] Samaras, G. M. and R. L. Horst (2005). "A systems engineering perspective on the human-centered design of health information systems." *Journal of Biomedical Informatics* 38(1): 61-74.
- [Seffah05] Seffah, A., M. C. Desmarais, et al. (2005). *HCI, Usability and Software Engineering Integration: present and future*. Human-Centered Software Engineering - Integrating Usability in the Software Development Lifecycle. A. Seffah, J. Gulliksen and M. C. Desmarais. Netherlands, Springer. 8: 37-57.
- [Seffah05a] Seffah, A., J. Gulliksen, et al. (2005). *An Introduction to Human-Centered Software Engineering: Integrating Usability in the Development Process*. Human-Centered Software Engineering — Integrating Usability in the Software Development Lifecycle, Springer Netherlands. 8: 3-14.
- [Sommerville07] Sommerville, I. (2007). *Software Engineering*. Harlow, Addison-Wesley.
- [Staccini01] Staccini, P., M. Joubert, et al. (2001). "Modeling health care processes for eliciting user requirements: a way to link a quality paradigm and clinical information system design." *International Journal of Medical Informatics* 64(2-3): 129-142.
- [Sutcliffe05] Sutcliffe, A. G. (2005). *Convergence or Competition between Software Engineering and Human Computer Interaction*. Human-Centered Software Engineering - Integrating Usability in the Software Development Lifecycle. A. Seffah, J. Gulliksen and M. C. Desmarais. Netherlands, Springer. 8: 71-84.
- [Teixeira08] Teixeira, L. (2008). *Contribuições para o Desenvolvimento de Sistemas de Informação na Saúde: Aplicação na Área da Hemofilia* (PhD Thesis - In Portuguese). Department of Economics, Management and Industrial Engineering, University of Aveiro. Aveiro, Aveiro University.
- [Zhang05] Zhang, J. (2005). "Human-centered computing in health information systems Part 1: analysis and design." *Journal of Biomedical Informatics* 38(1): 1-3.
- [Zhang02] Zhang, J., V. L. Patel, et al. (2002). "Designing human-centered distributed information systems." *IEEE Intelligent Systems* 17(5): 42-47.
- [Zimmermann07] Zimmermann, D. and L. Grötzbach (2007). *A Requirement Engineering Approach to User Centered Design*. Human-Computer Interaction. Interaction Design and Usability: 360-369.



# GUI Behavior from Source Code Analysis

João C. Silva<sup>1,2</sup>

Carlos E. Silva<sup>1</sup>

José C. Campos<sup>1</sup>

João A. Saraiva<sup>1</sup>

<sup>1</sup> Departamento de Informática/CCTC, Universidade do Minho

<sup>2</sup> Departamento de Tecnologia, Instituto Politécnico do Cávado e do Ave

{jose.campos, jas}@di.uminho.pt, jcsilva@ipca.pt, carlosebms@gmail.com

---

## Abstract

*When developing interactive applications, considering the correctness of graphical user interfaces (GUIs) code is essential. GUIs are critical components of today's software, and contemporary software tools do not provide enough support for ensuring GUIs' code quality. GUIsurfer, a GUI reverse engineering tool, enables evaluation of behavioral properties of user interfaces. It performs static analysis of GUI code, generating state machines that can help in the evaluation of interactive applications. This paper describes the design, software architecture, and the use of GUIsurfer through an example. The tool is easily re-targetable, and support is available to Java/Swing, and WxHaskell. The paper sets the ground for a generalization effort to consider rich internet applications. It explores the GWT web applications' user interface programming toolkit.*

## Keywords

*Graphical user interface, Reverse Engineering, Analysis*

---

## 1 Introduction

Practice shows that the user interface layer of interactive applications is the one more likely to suffer changes during the life-time of an application. Available technology to build user interfaces mostly consists of libraries of components that are glued together in an event-based style of programming, leading to code that is hard to understand and maintain. For example, the fact that Swing [Walrath 04] components are based on the Model-View-Controller (MVC) architectural pattern, does not necessarily mean that the pattern is maintained at application level. Indeed, the source code of Swing-based user interfaces can quickly become a collection of method calls accessing different parts of some common global state, reviving the notion of “spaghetti code”. The multitude of technologies and frameworks being made available for developing Web applications, if anything, is making things worst [Mikkonen 07].

Integrated Development Environments (IDEs) help developers in creating the user interfaces by allowing them to *draw* the user interface, and attach methods to relevant object/events. However, they do not necessarily promote better code structuring and quality.

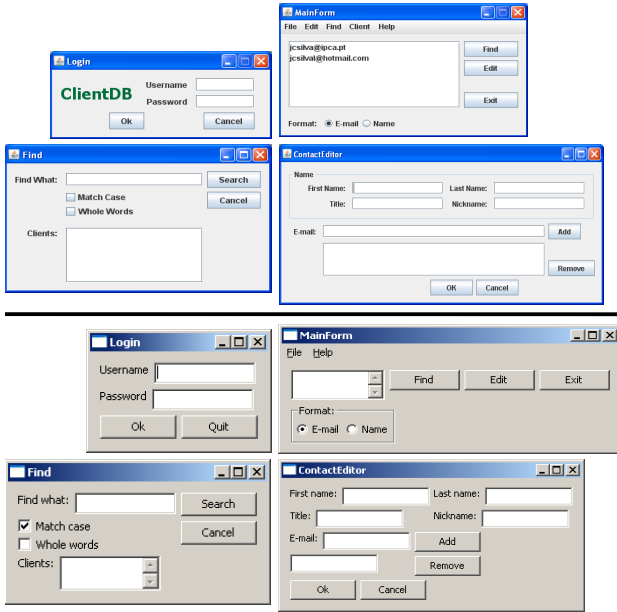
The above two issues, the need to constantly change code, and poor code quality from the start, mean that a considerable effort needs to be invested into the development and maintenance of the GUI layer of applications. Tools are needed that can help in this process. Reverse engineering tools, in particular, can have a role in helping analyse, un-

derstand, and manipulate source code.

Our objective consists in developing tools to automatically extract models from GUI source code. The extracted models should specify, for example, when a particular GUI event can occur, which are the related conditions, which system actions are executed or which GUI state is generated next. We want to be able to reason about, and test, this GUI model in order to analyze aspects of the original application's usability, and the quality of the implementation. Additionally, we want the developed tools to be language independent. Through the use of generic techniques, the tool enable to analyze different source code paradigms, such object oriented or functional. This work will not only be useful to enable the analysis of existing interactive applications, but can also be helpful when an existing application must be ported or simply updated [Melody 96].

In previous papers [Silva 06, Silva 09] we have explored the applicability of slicing techniques to our reverse engineering needs, and developed the building blocks for the approach. In this paper we describe our tool making use of two Agenda interactive application as running examples. A new module is presented, allowing us to analyze GWT-based rich internet applications.

The paper is organised as follows: section 2 describes the running example; section 3 describes the reverse engineering approach; section 4 describes some of the models the tool is able to generate; section 5 discusses analysis; section 6 discusses rich internet applications' support; section 8 discusses related work; and section 9 concludes with



**Figure 1. Two agenda applications — Java/Swing (top) and WxHaskell (bottom)**

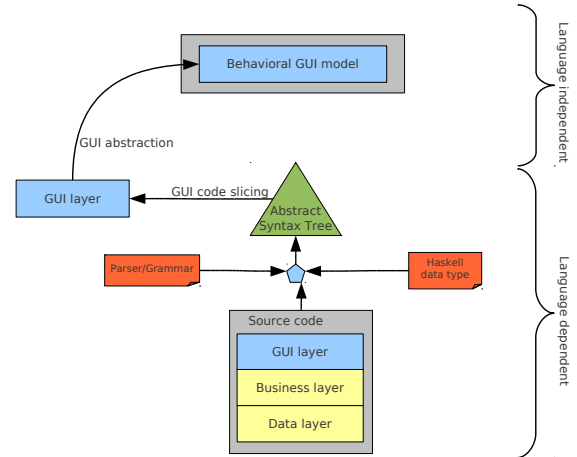
some pointers to future work.

## 2 The Agenda Example

Throughout the paper we will use two Agenda interactive applications as running examples. The first one was implemented with Java/Swing [Loy 02]. The second in WxHaskell [Jones 99]. These applications implement an agenda of contacts: they allow users to perform the usual actions of adding, removing and editing contacts. Furthermore, they also allow users to find a contact through its name.

Each agenda consists of four windows, named *Login*, *MainForm*, *Find* and *ContactEditor*, as shown in Figure 1. The initial *Login* window (Figure 1, top-left window) is used to control users' access to the agenda. Thus, a login and password have to be introduced by the user. If the user introduces a valid login/password and presses the *Ok* button, then the login window closes and the main window of the application is displayed. On the contrary, if the user introduces an invalid login/password, then the input fields are cleared, a warning message is produced and the login window continues to be displayed. By pressing the *Cancel* button in the *Login* window, the user exits the application.

Authorized users can use the main window (Figure 1, top-right window) to find and edit contacts (*Find* and *Edit* buttons). By pressing the *Find* button in the main window, the user opens the *Find* window (Figure 1, bottom-left window). This window is used to search and obtain a particular contact's data from his name. By pressing the *Edit* button in the main window, the user opens the *ContactEditor* window (Figure 1, bottom-right window). This last window allows the edition of all contact data, such as name, nickname, e-mails, etc. The *Add* and *Remove* buttons en-



**Figure 2. Model-based GUI reasoning process**

able edition of the e-mail addresses list of the contact. If there are no e-mails in the list then the *Remove* button is automatically disabled.

## 3 GUI Models from Source Code

In order to achieve our goal of developing an approach for reverse engineering of GUI source code, we resorted to several techniques. Figure 2 describes our approach.

Using a parser, an Abstract Syntax Tree (AST) is obtained from the source code. Then we identify all fragments in the abstract syntax tree that are members of the GUI layer. This is achieved through program slicing [Tip 95]. We use the GUI constructors to focus the slicing in the subtrees that represent the GUI. Slicing is based on the program dependency graph.

The *GUI code slicing* module extracts graphical user interface AST fragments through code slicing. This is a generic module to extract GUI fragments from any AST, i.e. Java/Swing, wxHaskell, C#, etc. This allows us to identify all of the program fragments that interact with the graphical user interface. We do a traversal of the tree (based on the program dependency graph), and detect all GUI nodes.

In order to extract the user interface behavior from the source code of the interactive applications, we need to construct a slicing function that isolates a sub-program from the entire program. Because we want to reuse our approach across different programming languages and paradigms, we need to use generic techniques that work for *any* AST and not for a particular language only.

Using strategic programming [Visser 03, Visser 04] we make use of a pre-defined set of (strategic) generic traversal functions that traverse any AST using different traversal strategies (e.g. top-down, left-to-right, etc). Thus, the programmer has to focus in the nodes of interest, only. In fact, the programmer does not need to have a knowl-

edge of the entire grammars/AST, but only of those parts he is interested in (e.g., the Swing sub-language). As a result, he/she does not need full knowledge of the grammar to write recursive functions that isolate the graphical user interface sub-program from the entire program. We used *Haskell* to develop a GUI code slicing library which contains a generic set of traversal functions that traverse any AST.

Figure 2 shows also that the *GUI abstraction* module uses the GUI fragments to produce a behavioral user interface description (the *Behavioral GUI model*). The fragments relevant to the GUI reverse engineering are limited to graphical user interface instructions, control flow information, and methods invocation. From these fragments of the original AST it is finally possible to extract the GUI layer and reason about it.

The implemented prototype is language independent in what regards the strategic programming and slicing techniques. The prototype is language dependent with respect to the grammar and program dependency graph definition.

#### 4 GUI Models

The abstractions that we look for in the source code are widgets that enable users to input data (*user input*), widgets that enable users to choose between several different options such as a command menu (*user selection*), any actions that are performed as the result of user input or user selection (*user action*), and any widget that enables communication from application to users such as a user dialogue (*output to user*). Given the user interface code of an interactive system and this set of abstractions, we can generate its graphical user interface abstraction.

Currently, the tool is capable of automatically generating models of the interface. Next we describe examples that we can automatically generate from the Agenda applications' source code.

The first example, presented in Figure 3, is a directed graph describing the Agenda application behavior. Interactive systems can be represented as directed graphs [Melody 96]. User actions are mapped into arcs and states are GUI application idle time. When the user performs an action, the current state A is changed to the next state B where there is a directed arc from A to B labeled with that action. Arcs may point back to the same state, and the transition then does not change the state. Graph models may be non-deterministic because of the underlying system, in which case one of several possible next states will be arrived at.

The directed graph presented in Figure 3 is useful to visualize application's windows states. Each state of this model contains the description of all windows opened in a particular period of time. Transitions between states correspond to events that allow to open or close windows. Each transition refers first the window's name, its state status, the event and respective condition. In this case, we can reason about which windows can be opened along a session, which are the related events, and which conditions must hold. As

an example, at the top left corner, this model specifies the Login window as the entry point for the application. Then, from the Login window, there is one transition to the MainForm window through the Ok event and cond2 condition pair (each event and condition identifiers are related to the respective source code of the original application). The referred transition is:

*Login state1 Edit cond2*

The state1 identifier refers to the internal state of the Login window. Although not illustrated in this paper, it is possible to generate models representing how the internal state of a windows changes. From MainForm, it is possible, for example, to open the ContactEditor window through the Edit/cond2 pair.

In this particular case the Java and Haskell Agenda applications have the same behavior. Hence, the graph from Figure 3 can be obtained from both. As an example, let us consider the following Java/Swing and WxHaskell source code which opens a new ContactEditor window:

```
...
private void EditActionPerformed(...) {
    new ContactEditor().setVisible(true);
    ...
}
```

and

```
...
edit <- button pn [
    text := "Edit",
    on command := start contactEditor
]
...
```

This source code is abstracted to the states in Figure 3 that make reference to the ContactEditor form.

Another example of a model generated by GUIsurfer is presented in Figure 4. This model contains all possible states for each application window. The model presents also the total number of events associated to each state transition. This is useful as a metric to detect windows complexity.

These examples provide an indication of how the developed prototype achieves our main objective. It generates, automatically, a GUI behavioral model directly from an application's source code. The behavioral model describes window states, events, conditions and actions.

We now want to be able to reason about this GUI model. The techniques described in this paper enable us to analyse properties of the interface. For example, we can use graph-based algorithms to compute if all states are accessible from the initial one, in order to detect whether a particular window of the application will ever be displayed or not. We can also produce valid or invalid *sentences* of the language defined by the machine to use as test cases. These test cases can be used to prove more advanced properties of the interface. The next section will show how we can reason about the original application's characteristics.

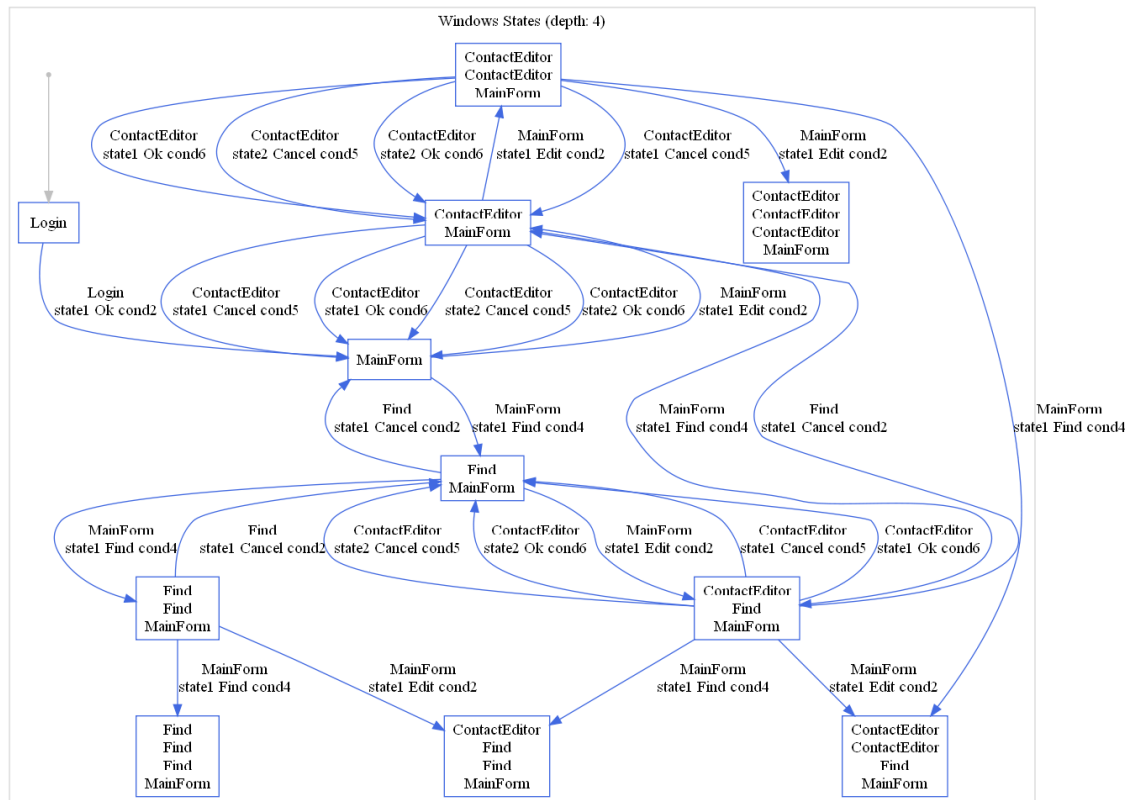


Figure 3. Agenda GUI state machine

## 5 Analysis

The reverse engineering approach described in this paper allows us to extract an abstract GUI behavior specification. Now, GUI analysis can be performed.

Graphs are a mathematical concept that can be used to model Graphical User Interfaces [Thimbleby 08]. Sequences of user actions are paths in a graph. A standard graph theoretic concept is the shortest path between two edges, which defines the most efficient way a user can achieve a particular change of state. Graphs define interactive systems and usability properties.

Using graph models we extended our prototype implementing graph operators. At this time we have implemented intersection, union and difference of graphs. This is particularly useful to compare versions of an application, allowing to determine different versions have the same behavior.

Let us consider a new version of the Java/Swing Agenda application without the Contact Editor form. I.e., without the following Java/Swing instruction.

```
new ContactEditor().setVisible(true);
```

Using the difference graph operator, we are able to obtain behavioral differences between applications. For example, calculating the difference between the WxHaskell and Java/Swing versions of the Agenda, we obtain the graph in Figure 5. The obtained graph identifies the actions which

can only be executed in the WxHaskell Agenda implementation.

Additionally, we can make use of the *QuickCheck Haskell* library tool. QuickCheck [Claessen 00] is a tool for testing *Haskell* programs automatically. QuickCheck provides combinators to define properties, observe the distribution of test data, and define test data generators. GUIsurfer is capable of generating a *Haskell* program capturing the behaviour of the application. Then properties may be defined and QuickCheck may be used to test them in a large number of randomly generated cases.

## 6 Rich Internet Applications

RIAs are an emergent technology whose primary goal is to develop web applications with the strengths of desktop applications. The principal advantages of desktop applications in comparison to traditional web applications are [Silva 10]:

- absence of page reloading;
- no need for an on-line connection;
- easy interaction with other desktop applications;
- superior interaction experience and usability;

However, traditional web applications, applications accessed through the network(internet/intranet), also have specific advantages such as:

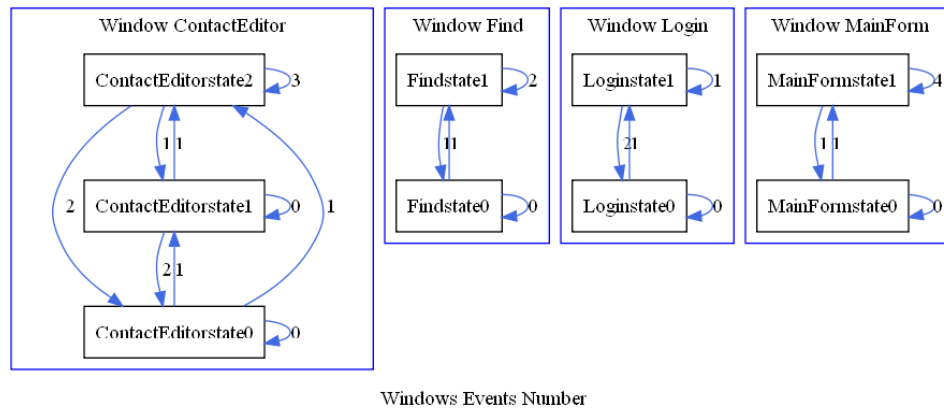


Figure 4. GUI state events number

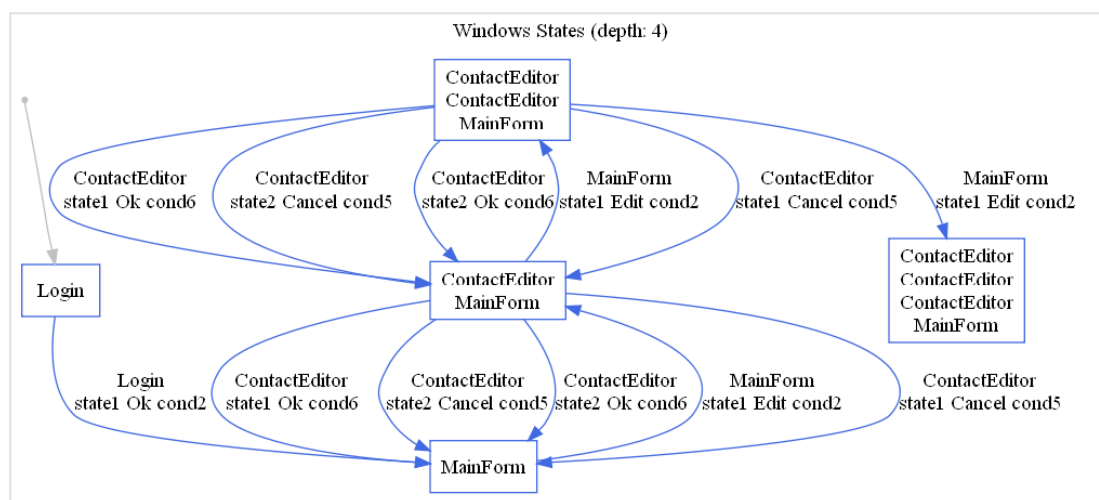


Figure 5. Comparing version of an application (using difference operator)

- they do not require deployment/installation or updates in every desktop;
- they provide easier access since one just needs an internet connection;
- they are easily made available in more platforms;
- concentrated information eases security and backups processes;

RIA technologies attempt to bring together these two worlds.

The question, then, is whether GUIsurfer can be extended to deal with this new and emerging programming paradigm.

### 6.1 Google Web Toolkit

Google Web Toolkit (GWT) is a technology that allows for the development of rich internet applications using the Java programming language. Since GUIsurfer is already able to reverse engineer Java/Swing programs, extending it to deal with GWT seemed like a logical next step.

As it can be deduced by its name, GWT was developed and introduced by Google (version 1.0 was released in May 2006). GWT is a set of development tools, programming utilities, and widgets that enable the user to create rich internet applications. Developing the application in the Java language allows GWT to bring all of Java's benefits to RIAs. Consequently, GWT aims to make the coding of RIAs as simple as possible while allowing interaction with existing JavaScript code.

GWT's goal is to make it easy to develop complex cross-browser applications. To this end, GWT provides a set of ready-to-use user interface widgets that can be immediately utilized to create new applications. Moreover, it also provides a simple way to create original widgets by combining the existing ones. Since GWT produced a JavaScript application, it does not require browser plug-ins additions, and there is also no need for an application server if the applications run fully on the client-side.

By making the development in the Java programming language, GWT inherits many of Java's benefits. One of these benefits is that it provides better application management

(when compared to plain JavaScript), thus making GWT a proper solution for the development of Web applications with significant size. This occurs as Java is an object oriented language, therefore allowing Java projects to generally be easy to communicate and comprehend.

Another advantage of using the Java language arises as it enables using any Java Integrated Development Environment (IDE) during application development. A Java IDE improves development as they provide several tools to help developers, for instance, code completion or error checking, and even tools to help debugging the application. Moreover, by using the Java language, one benefits also from Java type checking, therefore decreasing the number of application errors. There is also an improvement on JavaScript debugging, as errors are noticed in compilation time instead of execution time.

## 6.2 GWT support in GUIsurfer

RIAs have a far greater division between the interface code and the rest of the code, since they divide the code into client-side and server-side. This division is important since it helps GUIsurfer perform GUI code slicing, as the code is more partitioned. GUIsurfer possesses an architecture with a well defined distinction between the language dependent modules and language independent modules. The goal of generalizing it to GWT is made easier because there's only the need of restructuring the language dependent modules.

Since GWT it is a Java toolkit, the same parser already used by GUIsurfer for Java/Swing code could be used. Ideally then there would only be the need to perform the slicing step with a different set of GUI components (those of GWT instead of those from Swing). However a few issues arose. The first related to the genericity of the tool and was due to GUIsurfer's original implementation using the *addActionListener* method of Swing components to identify actions. In GWT however methods are registered through the *addClickHandler* method. Solving this problem meant parameterizing GUIsurfer on the method used to register event handler in the interface.

A second issue related to differences in the functionality of both toolkits (Swing and GWT). Since a GWT application is a web application, the closing window (in GWT, panels) actions available in Java Swing are not present. Closing a web application is an unusual action, and thus there is no direct support in GWT for doing it, though it can be achieved by invoking native JavaScript. Another issue occurred in detecting a change from a window/panel to another. In Swing this is achieved by invoking the *dispose* method on a class. In GWT this is achieved by making the visibility attribute of the panels. Again, changes were introduced to address this situation.

Once the above issues were addressed we were able to generate similar models to those on figure 1 for a GWT version of the Agenda application. In this first version of GWT support, an assumption is also made that the GWT code is structured as similar as possible to Java Swing code. Work is currently ongoing and our goal is to loosen these restric-

tions as much as possible, and generally improve support for panel handling.

## 7 A Language Independent Tool

In this section the applicability of GUIsurfer to GWT and WxHaskell code is discussed. Our retargets to Wx/Haskell and GWT highlight successes and problems with our initial approach. The size of the adaptations, and the time it took to code them are distinct.

The adaptation to GWT was easier because it exploited the same Java parser. The adaptation to Wx/Haskell was more complex as not only the language was different and the same parser could not be used, but also the programming paradigm was different, i.e. Haskell is a functional language. The functional paradigm is a programming paradigm that treats computation as the evaluation of mathematical functions and avoids state. For the applicability of GUIsurfer to Wx/Haskell we implemented the slicing step for functional programming analysis extracting events, related conditions and GUI actions through WxHaskell syntax. This task is more complex since Wx/Haskell toolkit has a different structure to define GUI components like windows, event actions, etc.

Regarding the applicability of GUIsurfer to GWT, we performed the slicing step with the set of GUI components from GWT, which are different those from Swing. Additional structures are different and need slicing adaptation such the *addActionListener* method of Swing components to identify actions. In GWT the respective method is registered through the *addClickHandler* method. Changes performed to extend GUIsurfer to a new programming language, specifically GWT or WxHaskell, didn't reflect on architectural alterations. Hence, GUIsurfer's objective of being a re-targetable tool was accomplished.

## 8 Related Work

We have described an approach to the analysis of interactive systems from source code. Reverse engineering techniques are used to derive models from the source code of the user interface layer. State machines are used to capture the behaviour of the interface. Graph theory is then used to analyse those state machines.

Having described the approach, we now set it in the general context of current approaches to the reverse engineering, modelling and analysis of interactive systems.

### 8.1 Reverse engineering

A typical reverse engineering approach is to run the interactive system and automatically record its state and events. Memon et al. [Memon 03] describe an tool which automatically transverses a user interface in order to extract information about its widgets, properties and values. Chen et al. [Chen 01] propose a specification-based technique to test user interfaces. Users graphically manipulate test specifications represented by finite state machines which are obtained from running the system. Systa studies and analyses the run-time behaviour of Java software through a

reverse engineering process [Systa 01]. Running the target software under a debugger allows for the generation of state diagrams.

Another alternative is the use of static analysis. The reengineering process is based on analysis of the application's code, instead of its execution, as in previous approaches. One such approach is the work by d'Ausbourg et al. [d'Ausbourg 96] in reverse engineering UIL code (User Interface Language – a language to describe user interfaces for the X11 Windowing System, see [Heller 94]). In this case models are created at the level of the events that can happen in the components of the user interface. For example, pressing or releasing a button.

Moore [Moore 96] describes a technique to partially automate reverse engineering character based user interfaces of legacy applications. The result of this process is a model for user interface understanding and migration. The work shows that a language-independent set of rules can be used to detect interactive components from legacy code. Merlo [Merlo 95] proposes a similar approach. In both cases static analysis is used.

We are using static analysis as in [Merlo 95, Moore 96, d'Ausbourg 96]. However, we are applying it to the source code of graphical user interfaces developed in general purpose programming languages, and working on making the approach as language independent as possible.

## 8.2 Modelling and analysis

State machines and graph theory are common in the modelling and analysis of interactive systems. Horrocks presents a proven technique for designing event-driven software using the statechart notation [Horrocks 99]. With statecharts it is possible to model multiple cross-functional state diagrams within the statechart. Each of these cross-functional state machines can transition internally without affecting the other state machines in the statechart.

Thimbleby [Thimbleby 08] gives examples of the use of graph theory in the modelling and analysis of a real interactive device. The work described a variety of graph theoretic properties, and discuss their significance to interaction design. Graph theory was also proposed for use in human computer interaction in [Memon 01] as a means of analysis. A representation of a GUI component, called an event-flow graph, identifies the interaction of events within a component.

Other work includes using graph theory for providing test models [Lu 08, Li 07]. Automated graphic user interface test models, which are based on the event-flow graph, are proposed.

Our work builds on these approaches to define appropriate models and analysis approaches to be supported by the GUIsurfer tool.

## 9 Conclusions

When developing interactive applications, considering the correctness of the graphical user interface code is essential. GUIs are critical components of today's software and

contemporary software tools do not provide enough support for guaranteeing GUI code quality and maintenance. With this in mind, we are developing GUIsurfer, a tool to reverse engineer the GUI layer of interactive applications.

This work is an approach for improving analysis techniques allowing us to reason about GUI models through graph theory. We described GUI models extracted automatically from source code, and presented a methodology to reason about the user interface model.

The approach is language-independent. We have applied the techniques to extract similar models from *Haskell/WxHaskell*, *Java/Swing* and *GWT* interactive applications. These models enable us to reason about both metrics of the design, and the quality of the implementation of that design. It is not possible with the actual prototype to analyze every existing JAVA/Haskell code. The system assumes source code will be structured according to specific conventions. In this case, it is assumed the code is generated with the conventions used by the NetBeans or WxHaskell integrated development environments (IDEs). To consider other kind of source code structuring, some adjustments would need to be made. Given that IDEs such as Netbeans are widely used, and automatically generate most of the user interface code, we do not believe this to be a major restriction.

Our objective has been to investigate the feasibility of the approach. We now plan to expand the tool at two levels. On the one hand, we will work on improving the GUIsurfer's support to the above mentioned programming languages, and extend it to new languages (for example, directly supporting the analysis of JavaScript code).

On the other hand, we plan to work on incrementing GUIsurfer's models generation capabilities to allow new types of analysis. Currently it produces state models, a type of dialog models. We want to investigate the feasibility of generating task models because they would allow for a more user centered evaluation of the system's design.

## Acknowledgements

GUIsurfer development is being carried out in the context of the CROSS and SSaaPPP projects, supported by the Portuguese Research Foundation (FCT) under contracts: PTDC/EIA-CCO/108995/2008 and PTDC/EIA-CCO/108613/2008.

João C. Silva is supported by PhD grant SFRH/BD/30729-/2006 from FCT.

## References

- [Chen 01] J. Chen e S. Subramaniam. A gui environment for testing gui-based applications in java. *Proceedings of the 34th Hawaii International Conferences on System Sciences*, 2001.
- [Claessen 00] Koen Claessen e John Hughes. Quickcheck: A lightweight tool for random testing of haskell programs. Em *ICFP, ACM SIGPLAN, 2000*, 2000.



- [d'Ausbourg 96] Bruno d'Ausbourg, Guy Durrieu, e Pierre Roché. Deriving a formal model of an interactive system from its UIL description in order to verify and to test its behaviour. Em *DSV-IS 96*. 1996.
- [Heller 94] Dan Heller e Paula M. Ferguson. *Motif Programming Manual*, volume 6A de *X Window System Seris*. O'Reilly & Associates, Inc., second edição, 1994.
- [Horrocks 99] Ian Horrocks. *Constructing the User Interface with Statecharts*. Addison-Wesley, Harlow, England, 1999.
- [Jones 99] Simon Peyton Jones, John Hughes, Lennart Augustsson, et al. Report on the programming language haskell 98. Relatório técnico, Yale University, Fevereiro 1999.
- [Li 07] Ping Li, Toan Huynh, Marek Reformat, e James Miller. A practical approach to testing gui systems. *Empirical Softw. Engg.*, 12(4):331–357, 2007.
- [Loy 02] Marc Loy, Robert Eckstein, Dave Wood, James Elliott, e Brian Cole. *Java Swing, 2nd Edition*. O Reilly, 2002.
- [Lu 08] Yongzhong Lu, Danping Yan, Songlin Nie, e Chun Wang. Development of an improved gui automation test system based on event-flow graph. Em *CSSE '08: Proceedings of the 2008 International Conference on Computer Science and Software Engineering*, páginas 712–715, Washington, DC, USA, 2008. IEEE Computer Society.
- [Melody 96] Moore Melody. A survey of representations for recovering user interface specifications for reengineering. Relatório técnico, Institute of Technology, Atlanta, 1996.
- [Memon 01] Atif M. Memon, Mary Lou Soffa, e Martha E. Pollack. Coverage criteria for gui testing. Em *ESEC/FSE-9: Proceedings of the 8th European software engineering conference held jointly with 9th ACM SIGSOFT international symposium on Foundations of software engineering*, páginas 256–267, New York, NY, USA, 2001. ACM.
- [Memon 03] Atif Memon, Ishan Banerjee, e Adithya Nagarajan. GUI ripping: Reverse engineering of graphical user interfaces for testing. Relatório técnico, University of Maryland, USA, 2003.
- [Merlo 95] E. Merlo, P. Y. Gagne, J.F. Girard, K. Kontogiannis, L.J. Hendren, P. Panangaden, e R. De Mori. Reengineering user interfaces. *IEEE Software*, 12(1), 64-73, 1995.
- [Mikkonen 07] Tommi Mikkonen e Antero Taivalsaari. Web applications - spaghetti code for the 21st century. Relatório Técnico SMLI TR-2007-166, Sun Labs, June 2007.
- [Moore 96] M. M. Moore. Rule-based detection for reverse engineering user interfaces. *Proceedings of the Third Working Conference on Reverse Engineering*, pages 42-8, Monterey, CA, november 1996.
- [Silva 06] J.C. Silva, José Creissac Campos, e João Saraiva. Combining formal methods and functional strategies regarding the reverse engineering of interactive applications. Em *DSV-IS 2006, Dublin, Ireland*. Springer, 2006.
- [Silva 09] J.C. Silva, José Creissac Campos, e João Saraiva. A generic library for gui reasoning and testing. *SAC ACM, Honolulu, USA*, March 2009.
- [Silva 10] Carlos Eduardo Silva. Reverse engineering of rich internet applications. Relatório técnico, Universidade do Minho, 2010.
- [Systa 01] T. Systa. Dynamic reverse engineering of java software. Relatório técnico, University of Tampere, Finland, 2001.
- [Thimbleby 08] Harold Thimbleby e Jeremy Gow. Applying graph theory to interaction design. páginas 501–519, 2008.
- [Tip 95] Frank Tip. A survey of program slicing techniques. *Journal of Programming Languages*, september 1995.
- [Visser 03] Eelco Visser. Program transformation with Stratego/XT: Rules, strategies, tools, and systems in StrategoXT-0.9. 2003.
- [Visser 04] Joost Visser e João Saraiva. Tutorial on strategic programming across programming paradigms. Em *8th Brazilian Symposium on Programming Languages*, Niteroi, Brazil, May 2004.
- [Walrath 04] Kathy Walrath, Mary Campione, Alison Huml, e Sharon Zakhour. *The JFC Swing Tutorial: A Guide to Constructing GUIs*. Prentice-Hall, 2nd edição, 2004.



# GUI for XML Documents Access using Query-By-Example Paradigm

Daniela Fonte Daniela da Cruz Pedro Rangel Henriques  
 Department of Informatics, 4710 Braga, Portugal  
 {danielafonte, danieladacruz, prh}@di.uminho.pt

Alda Lopes Gançarski  
 Institut Télécom, Télécom SudParis, CNRS UM Samovar  
 9 rue Charles Fourier, 91011 Evry, France  
 alda.gancarski@it-sudparis.eu

## Abstract

*In the context of our research project, we propose an interactive tool (GuessXQ) to perform the access to the information in a collection of eXtensible Markup Language (XML) documents. Due to the complex nature of those structured documents, the associated standard query language is also complex and, thus, not easy for most of the users. The GUI we propose does not require any knowledge about the query language, it is based on the Query-By-Example (QBE) paradigm from traditional databases. Using QBE, instead of specifying the desired components of the documents and eventual restrictions in a query, the user exemplifies those components marking them directly on a sample document picked-up from the collection. We believe this leverages the user's cognition about the object to search. GuessXQ is then responsible for the generation of the query to be treated by the information retrieval engine.*

## Keywords

XML, XQuery, Query-by-example.

## 1. Introduction

This paper addresses the problem of eXtensible Markup Language (XML) documents information access. Those documents, being structured, are accessed using specific query languages where the interesting structural components are specified, as well as restrictions over them if needed. The standard query language for XML is XQuery [BCF<sup>+</sup>05]. XQuery queries are powerful but complex to write (the user must have a deep knowledge of the query language as well as the document structure). To help the user in the task of specifying his queries, some specialized editors have been developed ([Kim02], [Oxy08]), but still requiring a good knowledge level of the query language.

At the same time, thinking that "Example is always more efficacious than precept", HCI researchers proposed a Query-by-example (QBE) paradigm as a new interaction mechanism in the context of database querying [RG07]. QBE is based on the concept that the user formulates his query by filling in the appropriate skeleton tables the fields (relational projection concept) and/or restrictions on fields (relational selection concept) he intends to search for.

Due to the complex nature of XML documents querying, the QBE concept was adapted to XML retrieval [LGB07], [BC05], [NO04] by showing the XML Schema Definition

(XSD) tree instead of the relational table skeleton. The system we are developing [dCFH<sup>+</sup>09], called GuessXQ, also displays the XML Schema tree representation. Moreover and distinct from other approaches, the user can query the entire collection exemplifying over a sample document. Thus, elements selection and restriction is done directly in the sample document, which gives the user a clear indication of the information he is searching for.

Suppose the user is willing to search in a ship order document composed by a set of ship destinations, where each one is defined by an address, a city and a country. Consider also that his specific interest is to look for the order identifier number 889923 and get all the *addresses* where the item with the title *The Secret* was shipped to. Adopting the QBE principle, instead of specifying the query in textual form, the user selects directly on the sample document exhibited by the QBE interface (see Fig. 1) the desired *id* (attribute value) and the elements *person* and *country*, just clicking and highlighting them.

In this paper, we fully describe the interface (actually a GUI) of GuessXQ that is responsible for all the interaction between the user and the system, providing to the user a simple but effective way of querying a collection of XML documents based on a *query-by-example* approach. Spe-

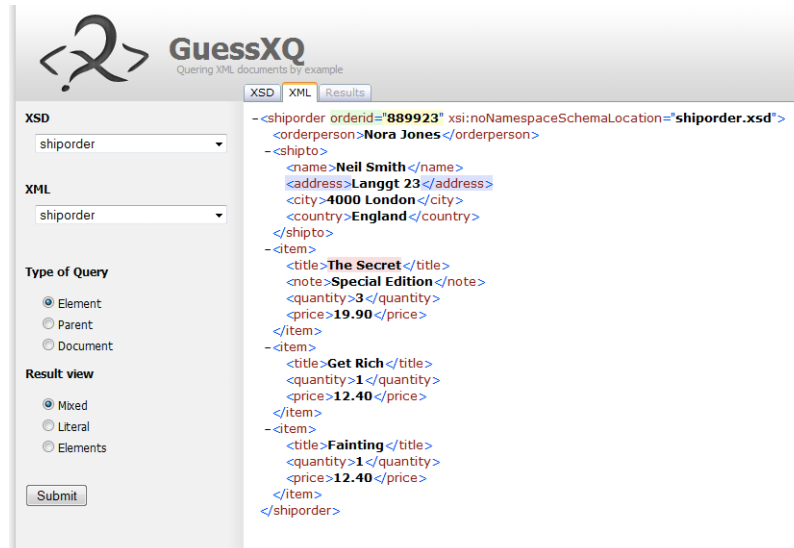


Figure 1. GuessXQ GUI (XML Tab): Sample document and Visual Query specification

cial care will be devoted to: the method used to choose the sample document from the entire collection; the visual specification of the query; the generation of an XQuery sentence from the visual specification; and the presentation of the documents retrieved.

The remainder of this paper is organized as follows. We first give an overview of the GuessXQ system discussing its architecture (Sec. 2). Sec. 3 presents the criteria used by our approach in order to select a sample document from the collection. Sec. 4 introduces the interface used to display the sample document and to allow for the visual specification of the query. This section elaborates on the way we implement the QBE paradigm. Sec. 5 describes the three different modes of queries offered by our system. Sec. 6 discusses the interface to display the information retrieved from the collection that satisfies the query. To conclude, in Sec. 7, we make some remarks and discuss the contribution of our approach, giving directions for future work.

## 2. GuessXQ System

After the choice of a XSD Schema (displayed in a specific window), representing the collection of documents from where the user wants to retrieve information, GuessXQ picks up a sample document from the collection and presents it (in a second window) to the user for him to specify his query. Fig. 1 shows the GUI where the user selects the components (elements or attributes) directly in the sample document. Another tab of that GUI is used to show the query inferred by the system and to display the documents retrieved.

After the *Sample Document Choice*, the document picked up is shown to the user in an interface for *Visual Query Specification*. This interface corresponds to the one presented above in Fig. 1. The *Visual Annotations* made in the previous interface are, then, translated into XQuery by a *Query Generator*. The generated query is processed by an *Information Retrieval Engine* which searches in the documents collection for the components specified in the

query. The returned components are given back to the user in a *Document Viewer* interface.

In the following paragraphs, we detail each bloc.

**XML and XSD Repository** The repository is a collection of XML files grouped by their schema (XSD). This Repository, in a simple way, is composed by two tables: *XMLdocs* and *XSDfamilies*. The *XMLdocs* table stores the name of the XML document, its location and its correspondent XSD family; the *XSDfamilies* table stores the name of each XSD and its location. To assess our system during development we are resorting to an archive of XML documents for testing, XAT (see [FCdC<sup>+</sup>10] for details) composed of documents belonging to the following Web accessible collections: Medical Subject Headings vocabulary files (MESH), the complete plays of Shakespeare, Eurovoc - a multilingual thesaurus covering the fields in which the European communities are active, and a set of miscellaneous documents collected from different sources.

**Collection Selection** The collection selection allows the other modules to access the repository in a systematic and simple way. It allows the other modules to select a Schema, a collection of documents, a single document or a component of a document. It also offers an appropriate interface to allow the user to choose a document type (XSD) from the repository.

**Sample Document Choice** The choice of the sample document is a crucial point in our approach, since the user specifies his needs over it. Thus, there must be a well founded logic behind the selection of the sample document from the collection. We have identified several criteria which should be taken into account when choosing that document (document size, number of different elements, diversity of values, among others), as discussed in Sec. 3.

**Visual Query Specification** The visual query specification is done through the sample document GUI (Fig. 1), which is responsible for the interaction between the user and the

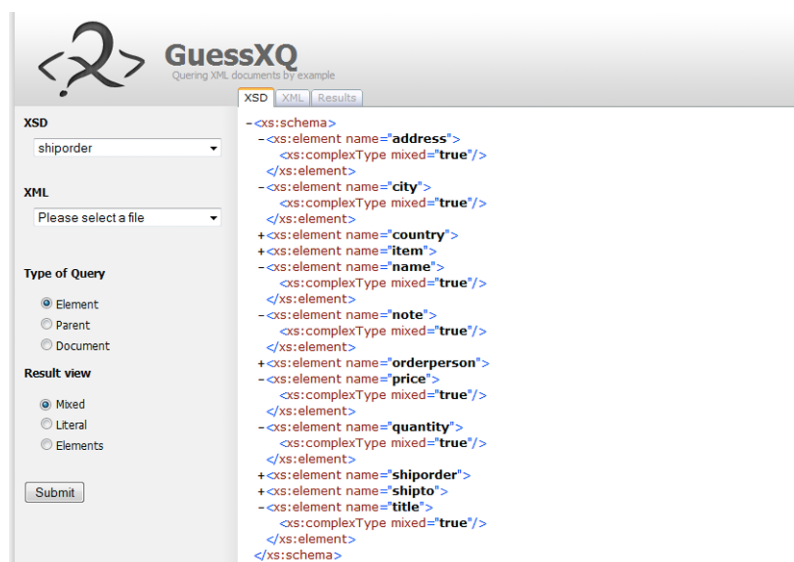


Figure 2. GuessXQ GUI (XSD Tab): the Collection Schema

system, in a way that allows the user to set the "example" for the QBE engine.

In our system, the specification of the query includes the selection of components (elements or attributes) and the possibility to restrict them to some desired value. Both of them are specified by clicking over the desired fields in the sample document. Each time a node is selected, its color is changed: this mechanism allows the user to easily see which nodes are already selected. It also uses different colors to distinguish between elements, attributes and values (restrictions) selection.

**Query Generator** After the query specification, the query generator module has the task of inferring the appropriate XQuery sentence. The selected nodes can be of two kinds: elements/attributes or values. An example of the former is selecting *person* element as shown in Fig. 1; the later happens for example when selecting the value from the attribute *id*. The construction of the query takes in consideration this two types of selected nodes.

**Information Retrieval Engine** After generating the query to retrieve document parts, a retrieval engine will access all documents in the collection where those parts appear. The engine has the capability to understand the query language and reach the interesting parts using several indexes.

**Retrieved Documents Viewer** The Retrieved Documents Viewer shows the results or answers produced by the retrieval engine. The user can choose, in the sample document interface (*Result view* option in Fig. 1), the type of the result to be shown: just elements; just textual values; or *Miscellaneous* (elements and text results in XML format). Fig. 4 is an example of the output visualization using *Miscellaneous* mode.

### 3. Sample Document Choice

In this section we will discuss the criteria used to rank documents. We identified the following metrics which should be taken into account when choosing the sample document.

**1st Document size** Big size files can slow down the system; but small size files can contain too little information or elements to aid the user selection. This metric can be used as a delimiter to complement the other ones; the principle is to not choose a file bigger than a predefined size.

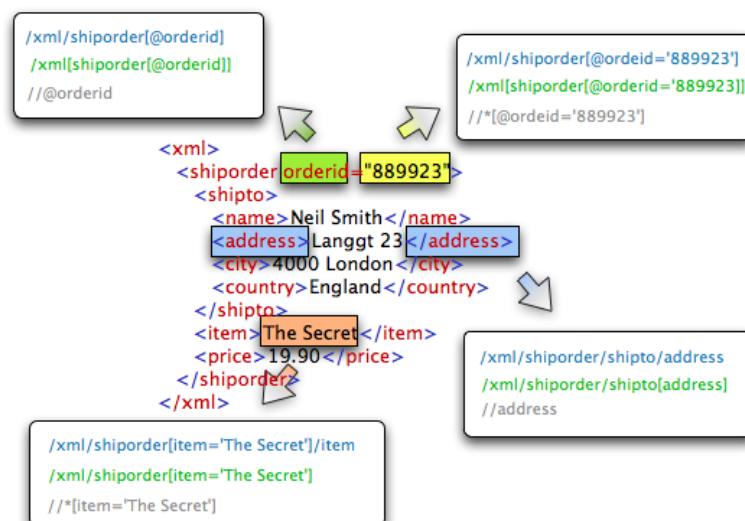
**2nd Number of components** It is important to take into account the number of components in the sample. On one hand, if the file has too many components, it can be too cluttered for the user to select the desired example. On the other hand, if the document has few components, maybe it does not contain all those the user needs.

**3th Number of different components** To counteract some of the shortcoming of the previous metric, it may be interesting to look at the number of different components in a file. This way, if a file contains almost all the elements and attributes present in the schema, the user gets a more complete variety of elements to specify his needs.

**4th Diversity of Values** As stated before, the main innovation of our QBE approach is the capacity of the user to see sample data and not just the structure (schema) of the queried documents. Therefore, a metric guaranteeing the diversity of data is important. Having different values for the same component allows the user to better understand the fields in the document he is querying. However, similar to the other metrics, if there is too much diversity, the example document may become too big.

**5th Number of commonly used components and values** It can be interesting to have a sample document as similar as possible to the majority of the documents in the collection. This means that the sample should contain commonly used components and values which may be the most interesting for the user example. A component/value is considered to be common if it is in the top N more frequent components/values of the collection.

As seen, each criterion has its own merits and shortcomings, so they must be used together in a meaningful way.



**Figure 3. Query Mode: the different XPath expressions associated to each component.**

#### 4. Visual Query Specification

GuessXQ offers a simple and intuitive GUI that allows the visual specification of the query, without the need of an advanced knowledge of the XQuery language. In this section, we describe the interaction between the system and the user through this interface (shown in Fig. 1, as referred in Sec. 2), to explain the QBE concept proposed.

As previously said, the user starts by selecting the intended collection choosing a Schema and the system shows its content in the XSD Tab, as illustrated in Fig. 2.

After the Schema selection **GuessXQ** suggests a sample document, as fully described in Sec. 3. To select the desired components (elements or attributes), the user can click directly on the specific item of the document (exhibited in the **XML** Tab, as illustrated in Fig. 1) and restricts their value. To improve this interaction, each time a component is selected its color is changed. This enhancement is obtained using different colors, according to their types: elements are highlighted in blue; attributes in green; **PCDATA** content is highlighted in red; and the attribute values in yellow. To unselect an highlighted component, the user just has to click again over it.

To improve the perception of the restriction implied by each selected item (in terms of the final query), when the user puts the mouse cursor over it, **GuessXQ** shows a tip with the correspondent XPath expression.

To simplify the search of the components to select, the system provides a feature to expand or retract blocks inside the document, more specifically elements with children nodes. Fig. 2 illustrates this feature: by clicking on ”-” (*minus*) sign, the correspondent block retracts, and only its element name is displayed, preceded now by a ”+” (*plus*) sign. To expand it again, the user just clicks on the ”+” sign, and the entire element subtree will be displayed.

After finish the visual query specification, the user must submit it so that **GuessXQ** starts the IR process, applying to the entire collection this query (see Sec. 6 for details).

## 5. Query Generator

GuessXQ generates the XQuery query based on the union of the paths obtained for each selected component. These paths are assigned to each document node according to three different modes of querying: *element-oriented*, *parent-oriented* and *document-oriented*, which user can choose in the *Type of view* option (as illustrated in Fig. 4). By default, system selects the *element-oriented* mode.

In the *element-oriented* mode, the path reflects the corresponding *absolute* location path from the root node to the selected component (as illustrated in Fig. 3 by the first path shown in each selection). In the *parent-oriented* mode, the generator assigns to each document node the *absolute* path from the root node to its parent (as illustrated in Fig. 3 by the second path in each selection). In the *document-oriented* mode, the generated XPath expression allows the search of the selected component on the entire document, independently of its position on the document tree (as illustrated in Fig.3 by the third path in each selection).

Before the submission, the user must select the type of view he wishes for the retrieved answer, as will be described in Sec. 6.

## 6. Retrieved Document Viewer

Concerning the visualization of the answer retrieved by the search engine, the user can choose one out of three modes (as said in Sec. 5): only the correspondent node names (*elements*); only the *literal* (PCDATA) values; or a *mixed* of both *elements* and *literal*, which shows the entire XML nodes resulting from the search. By default, the system outputs the results in *mixed* mode.

After the submission of the desired selection, **GuessXQ** shows in the **Results** Tab (as depicted in Fig. 4) the respective query and the result of applying it to the entire collection of documents (actually the desired output). First, the system displays the result extracted from the sample document, followed by the results (non-empty) extracted from each other document in the collection. Each result

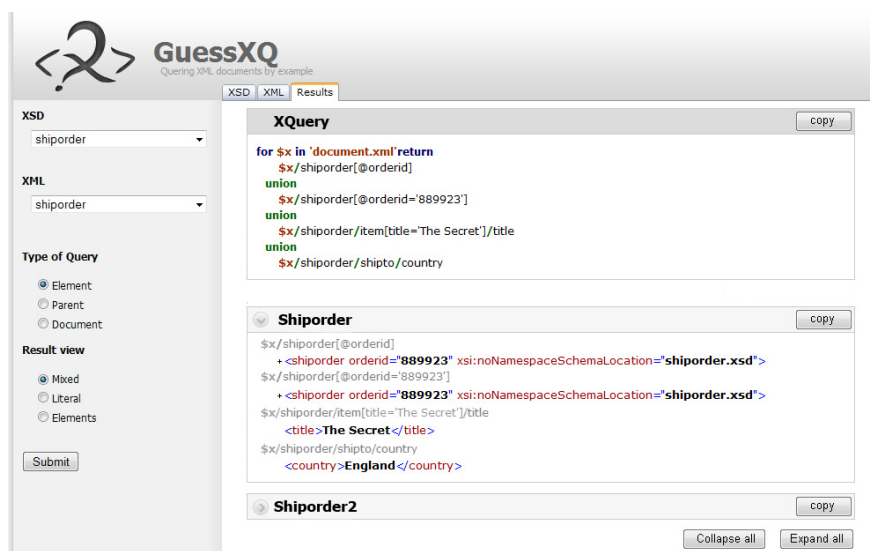


Figure 4. GuessXQ GUI (Results Tab): Retrieved Document Viewer

is loaded into a selector, to improve the usability: the user can expand the selector by clicking over it, or retracting it again by doing the same action (as illustrated in Fig. 4).

The system also allows to expand/collapse all the selectors with one click and offers the chance to copy the *query* or each particular output to the clipboard. In this way, we aim at offering a versatile and user-friendly interface.

## 7. Conclusion

This article presents new contributions for XML query specification in a user friendly interface. As far as we know, no other previous work addresses the framework of QBE in the context of *XML information access based on a sample document from the collection*.

In this paper, we focus on the general interface provided by our tool to support the visual specification of a query based on an example pointed out over a sample document. This interface also exhibits the query inferred from the selection, applies it to the entire collection, and at last displays the retrieved answer (a list of documents or their parts).

As future work, we plan to implement different methods to choose the sample document, as well as to finish other aspects under construction by now (like output visualization). After this, we think about making experiments for tuning the score computation parameters. As soon as possible, we intended to step forward to a second experimentation phase; GuessXQ interface will be assessed, testing the system with real users to measure their level of satisfiability.

## References

- [BC05] Daniele Braga and Alessandro Campi. Xqbe: A graphical environment to query xml data. *World Wide Web*, 8(3):287–316, 2005.
- [BCF<sup>+</sup>05] S. Boag, D. Chamberlin, M. Fernandez, D. Florescu, J. Robie, and J. Simeon. Xquery 1.0: An xml query language. w3c working draft. <http://www.w3c.org/TR/xquery>, 2005.
- [dCFH<sup>+</sup>09] Daniela da Cruz, Flávio Xavier Ferreira, Pedro Rangel Henriques, Alda Lopes Gançarski, and Bruno Defude. Guessxq, an inference web-engine for querying xml documents. In *INForum'09 — Simpósio de Informática*, pages 322 – 325, Lisboa, Portugal, September 2009. Faculdade de Ciências da Universidade de Lisboa.
- [FCdC<sup>+</sup>10] Daniela Fonte, Pedro Carvalho, Daniela da Cruz, Alda Lopes Gançarski, and Pedro Rangel Henriques. Xml archive for testing: a benchmark for guessxq. In *XATA 2010 — XML, Associated Technologies and Applications*, Vila do Conde, Portugal, May 2010.
- [Kim02] Larry Kim. *The XMLSPY Handbook*. John Wiley & Sons, Inc., New York, NY, USA, 2002.
- [LGB07] X. Li1, J. H. Gennari1, and J. F. Brinkley. Xgi: A graphical interface for xquery creation. In *Proceedings of the American Medical Informatics Association Annual Symposium*, pages 453–457. American Medical Informatics Association, 2007.
- [NO04] Scott Newman and Z. Meral Ozsoyoglu. A tree-structured query interface for querying semi-structured data. *Scientific and Statistical Database Management, International Conference on*, 0:127, 2004.
- [Oxy08] Oxygen xml editor. <http://www.oxygenxml.com>, 2008.
- [RG07] Raghu Ramakrishnan and Johannes Gehrke. *Database Management Systems*, chapter 6 - Query-by-Example (QBE). 2007.

Aveiro, 13, 14 e 15 de Outubro de 2010

## **Sessão IV**

### **Realidade Virtual e Aumentada**



Aveiro, 13, 14 e 15 de Outubro de 2010

# Crystal Ball, Virtual Gallery Creation System for Immersive Environments

Rui Nuno de Almeida  
MEI-M ISCTE-IUL - Lisboa  
rnaao@iscte.pt

Miguel Sales Dias  
ISCTE-IUL – Lisboa  
MLDC, Microsoft Language Development  
Center – Porto Salvo  
miguel.dias@microsoft.com

## Abstract

*This paper presents a concept and a tool suite for authoring, configuring, organizing and displaying 3D virtual galleries, depicting multimedia content, like art, picture, movie or any other kind of media, which are appropriate for large scale immersive virtual environment such as the CAVE™. A comparison was established by the authors with standard PC environments, while experimenting virtual environments produced by Crystal Ball, showing the superior experience obtained by an installation of the type of CAVE™ that ins operation at Lousal, south of Portugal.*

## Keywords

3D, Virtual Reality, Immersive Virtual Environments, On-line Virtual Worlds, Virtual Galleries

## 1. INTRODUCTION

Over the last years, 3D virtual environments and in particular, immersive 3d virtual environments, have turned into a popular tool specially for critical industries like architecture, oil exploration, aeronautics, flight simulation, defence training and medicine, amongst many others. These environments and new 3D user interfaces will likely provide the user with a different and more complete experience than the usual desktop PC with a keyboard, a mouse or a joystick. Content creation and application interfaces for such environments still have a long way to run, but it is now possible for people with limited knowledge of programming and/or scripting languages, to author, configure and use 3D and multimedia content for immersive environments.

To address the need of creating multimedia rich content to be experienced in large scale virtual environments, we propose a new concept and suite of tools, able to generate 3D and multimedia content for these, the Crystal Ball. The rich content produced by such tool suite, was tested both in desktop-based PC systems and in the CaveH, an immersive virtual reality environment that exists in Lousal, in south of Portugal. The system was developed in C++ and was based on the Openscenegraph [openscenegraph.org] and includes a series of tools available in the community. With Crystal Ball we can configure and customize galleries in a text file, which after being imported into the CaveH middleware will build a ready-to-use scene in the 3D virtual environment.

Crystal Ball galleries are made of several nodes. Each of these nodes may have a variable number of images, or

other “display elements”. Nodes can connect to other nodes by “link elements”, thus generating a structured network. All those nodes are defined in configuration text files, with the location and other parameters. Images and other media are also defined in text files that the system imports turning the settings into elements in the scene graph. These files can be constructed manually or by computer applications. Nodes can be independent pieces of geometry made by any 3D modeller or artist, in any software tool that produces the common 3D formats, such as 3DS or OBJ. The nodes of the gallery will have to comply with certain pre-determined size, so that the configuration of “displays” will stay in line with the respective node.

In this paper, we will start by providing in section 2, an insight in related work regarding virtual galleries authoring systems and tools. In section 3 we present the overall concept of how Crystal Ball defines virtual galleries. Section 4, presents the Crystal Ball scene description in more detail. In section 5, we present the basic technical approaches for the Crystal Ball content development and describe how we have performed PC and immersive virtual reality experiments, highlighting some of the findings. Finally, in section 6, we extract some conclusions and point some line of further research.

## 2. RELATED WORK

Three dimensional, spatially-oriented, exploratory and multiplayer 3D virtual environments are not a new thing, and have been used worldwide, particularly in education [Schendel08].

## 2.1 Virtual Worlds

In Virtual Worlds like Second Life® [SL09], Virtual Galleries, can be seen all over the Grid. “In 2007 there were more than 150 museums in Second Life, such as the International Spaceflight Museum (Spaceport Alpha 48, 78, 24), Second Louvre Museum (Tompson 153, 97, 100), Sci-Fi Museum (Indigo 75, 213, 22), Bayside Beach Galleria Museum of Contemporary Art (Flyingroc Chung 109, 118, 22), Paris 1900 (Paris 1900 9, 174, 16)” [Urban07]

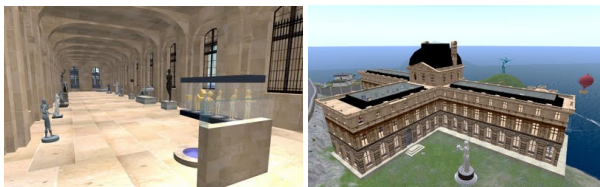


Figure 1– Outside and inside of Second Louvre Museum (Pictures: Bettina Tizzy).

Recently the makers of the 'Second Louvre Museum' (Fig. 1), announced that they intend to close its creation and have put on sale this replica of the real Paris museum. The Second Life world creation “bubble boom”, occurred specially during the year of 2007. At that time, thousands of companies created virtual environments to put their presence in this new platform for marketing and communications. The following year many companies have found that virtual worlds users (also known as “residents”), occupied themselves with the affairs of a social network, and rarely stuck to business premises and services. As a result, many of these companies and organizations have abandoned Second Life. Apparently, residents only occasionally looked for a virtual art gallery experience even when it is a spectacular creation such as the replica of the Louvre. The residents were in fact not interested and the number of visitors was very low. They looked mainly for other virtual spaces for social engagement. We can conclude that many of the residents want to create things for themselves and only occasionally visit big museums and artwork galleries.

Despite the proven disinterest of residents, organizations like ‘TheArchNetwork’ keep developing art projects like the recent 'Frank Lloyd Wright Virtual Museum' (Fig. 2), which opened during 2009 in Second Life®.



Figure 2 - Frank Lloyd Wright Virtual Is-land/Museum (Pictures: Palup Ling).

The site has now been officially recognized by the Frank Lloyd Wright Foundation, the Foundation created in 1940 by Frank Lloyd Wright and one of the greatest symbols of Modern Architecture.

On an island exclusively dedicated to Wright, visitors can depict many of its most emblematic works (mainly residential), as well as permanent and touring exhibits. The island / Museum have exhibitions like the one on 'Uson-

ian Houses' (the 'style' of American homes identified by Wright) [Frias09].

The most common exhibitions in Virtual Worlds are promoted and developed by the artists themselves (see Fig.3), to exploit their work and get close to the public. They usually use socially prestigious or crowded virtual spaces (hub Islands) to build the galleries.



Figure 3 - Exhibition of Vasco Lago Pinto (Jazz62 Masala in SL™). Black and White pictures of the artist exposed on “LX Gallery”, in Portucalis Island, in Second Life®. Galeria LX, on Portucalis the island, in Second Life®, part of 'Art Galleries in Second Life' network. A retrospective exhibition of all artists which exhibited their works over the year 2009.

## 2.2 CANVAS [Kaczmariski05]

A team at University of Illinois in USA, developed CANVAS (Collaborative Advanced Navigation Virtual Art Studio, Fig. 4). It is a room-sized immersive 3D environment with origins in CAVE™ technology from the Electronic Visualization Laboratory at the University of Illinois, Chicago. It is a scalable, reconfigurable display technology for modern art museums and is not intended as a work itself but as an environment to facilitate the creation and display of immersive art works. CANVAS is collaborative because it can be connected to an array of geographically dispersed immersive virtual spaces, has advanced navigation to allow viewers in different locations to interact with virtual art and allows for the creation and presentation of virtual art that exists not in two or three-dimensional space like a painting or sculpture, but in the multi-dimensional world of virtual images. Viewers see the CANVAS in 3D stereoscopy, by wearing passive stereo glasses.

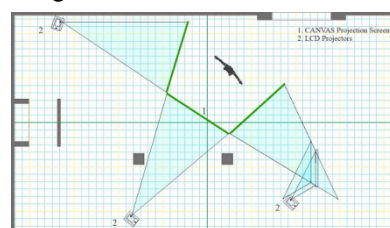


Figure 4 - The layout of the CANVAS at the Krannert Art Museum showing (1) the three-walled projection screen and (2) the set of six LCD projectors. Illustration by Camille Goudeseune.

KAMScript, [Marshack07] developed by the ISL and the Krannert Art Museum at the University of Illinois at Urbana-Champaign, is a system that allows non-programmer artists to create immersive 3D environments that can be experienced in the CANVAS.

## 2.3 VIRTUAL ART GALLERY TOOL [Semiao08]

Pedro Semião and Beatriz Carmo from the Science Department of Lisbon University, developed a Virtual Art Gallery tool set (Fig. 5). The purpose of this tool entirely

based in free software is to allow users to interactively create a virtual exhibition of artworks in a pre-built virtual model. It uses X3D for the models and Java with Xj3D for the display and handling. The works of art are handled through a MySQL database. The toolset comprises the Virtual Art Gallery tool, the Space Picker application and the Virtual Exhibition Builder application (Fig. 6).

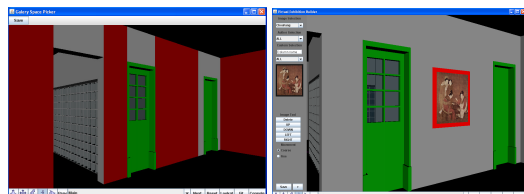


Figure 5 - The Space Picker application and Virtual Exhibition Builder application.

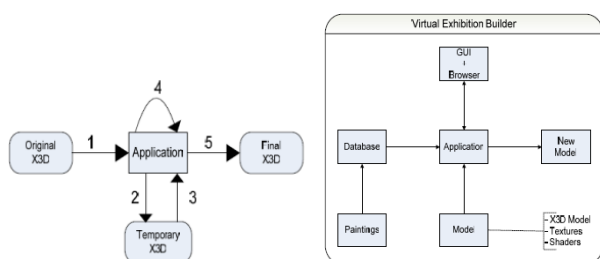


Figure 6 - Space Picker Data Flow and Virtual Exhibition Builder Application Layout.

## 2.4 XVM scene description [Lu08]

XVM [LU 2008], is a XML based scene description language for 3D virtual museum (Fig. 7). XVM describes several classes of nodes like dummy nodes, built-in shape nodes, model appearance nodes, line wall nodes.

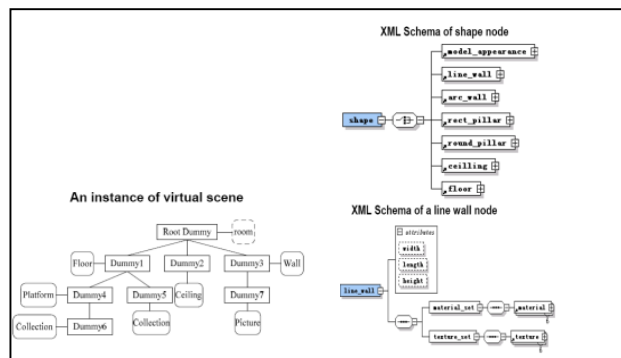


Figure 7 – Virtual scene and Node representations in XVM.

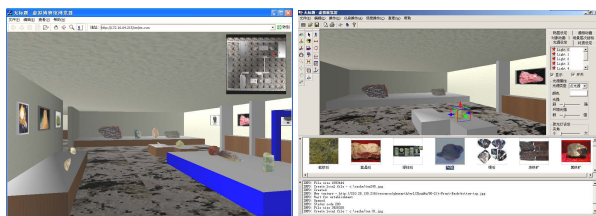


Figure 8 - GUI of virtual navigation and editing scene.

## 2.5 KAMScript describing [Marshack07]

Another way of describing a virtual gallery scene, is to use KAMScript (Fig. 9). This is a tool that allows people

with no programming experience to create three-dimensional virtual environments that can be displayed on the University of Illinois CANVAS system. Users simply create a text file that describes the location, orientation and size of 2D images (in JPEG file format) and 3D models (in WaveFront \*.OBJ format). KAMScript reads this text file, along with the supporting data files and renders the scene. With KAMScript, sounds (in the form of MP3 files) can be inserted into the virtual world at defined locations. Primitive volumes - such as boxes, ellipsoids, tori, cones and teapots – can be created in any RGB color. KAMScript essentially provides access to various OpenGL and Syzygy functions (the graphics and programming libraries used to create applications for the CANVAS cluster-based CAVE™), but does not require users to have full knowledge of what these functions are or how they are implemented.

```
// This is a comment.

IMAGE(helloworld.jpg,0,3,-2)

IMAGE(helloworld.jpg,[2,3,2],[0,45,0],[2,2])

IMAGE(helloworld.jpg,[-1,5,2],[45,0,0],[2,2])

SOUND(music.mp3,0,5,0)
```

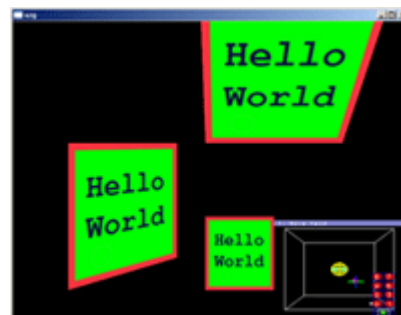


Figure 9 - An example of a KAMScript text file and resulting world. The first line places the image named “helloworld.jpg” at  $x=0$ ,  $y=3$ ,  $z=-2$ . The second line adds rotation and scaling as the second and third sets of parameters. The resulting image is shown using the standalone player.

## 3. CRYSTAL BALL (CBG)

We propose Crystal Ball as a concept and a series of tools to facilitate the authoring, configuring, organizing and displaying virtual galleries in 3D immersive virtual environments, which was much influenced by the CANVAS system. The Crystal Ball concept is based upon three entities – “nodes”, “links” and “displays”. To better understand what nodes and links are, we present a practical example of its parameters. We will start by assuming we want to create a room (node) for 6 planes “displays” with pictures or paintings of an artist. It is like a room in which the avatar will travel when visiting the Virtual Gallery. The designer and the gallery planner decide the basic dimensions for the nodes as well as the shape of the node, diameters and other measures as in the example of Fig 10.

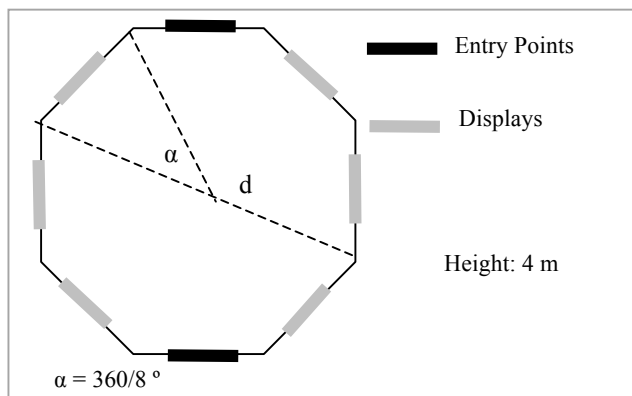


Figure 10 – Example of a node and his dimensions.

The modeller may create a *node* in accordance with the specifications and it will be added to the geometry of the scene. Each node can be reused as often as necessary in many other rooms or galleries. *Nodes* can be *empty* or *filled* (Fig. 11). A *filled node* is a node with a defined set of displays.

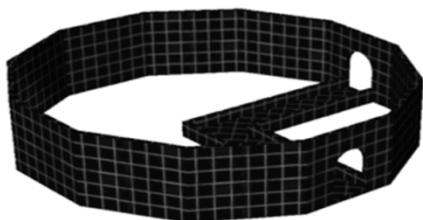


Figure 11 – Ready-made 3D empty node in Crystal Ball.

A *filled node* will have the displays placed on the defined geometric position. Nodes can have connections between them. Those connections link the *entry point* of one node to an *entry point* of another node. For this work we assume a maximum number of 4 links for each node. Seen from top view, the gallery will look like an organized grid structure that reminds a Printed Circuit Board (Figs. 12, 13).

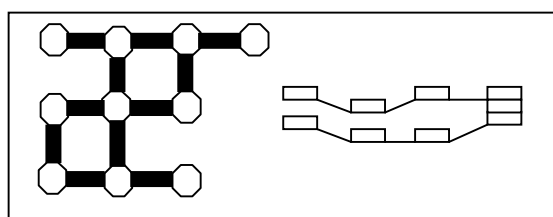


Figure 12 – Example of top view and side view of a CBG.

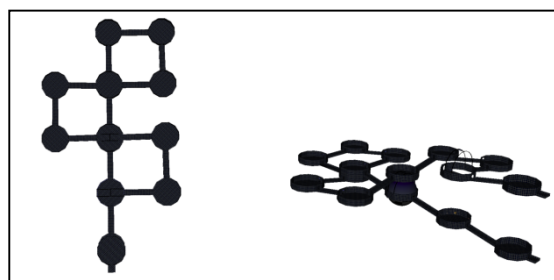
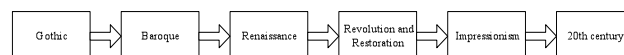


Figure 13 – Down view and perspective view of a CBG.

Each gallery can have an unlimited number of connected nodes. The only limit is machine storage and memory capacity. Based on a user-defined “flow file” and nodes topology, the system creates a corridor made of predefined planar or non-planar links.

Nodes can be connected according to a logic flow, in which the avatar will travel following a prescribed order that makes sense for that gallery in particular. The complexity of the structure depends on the context and multimedia contents of the gallery. The aim of the CBG system, is not the creation of art galleries simulating the real world ones, but rather create a logical navigation for the avatar in a tutorial manner. Suppose we want to create a gallery to make a virtual travelling system for the all contents of the Paris Louvre Museum. For sure we can imagine how big and complex that structure would be. We may want the avatar to flow inside the gallery with certain progression logic, like the following:



We must then build a linked structure between the nodes always following that order. This way we can define a workflow to orient and simplify the visiting or learning experience.

#### 4. SCENE DESCRIPTION IN CRYSTAL BALL

Our scene description based in text files, was the way to reach our goal of building the scene with the simplest structure, but is obviously extensible to something more structured, like the XML-based XVM scene description language for 3D virtual museums.

##### 4.1 Display Structure

In the Crystal Ball approach, we will simply call nodes to our display structures. Each node has a set of properties, and the gallery is the set of all nodes, links and displays.

##### 4.1.1 Plane Displays

Plane displays can be representations of pictures, paintings, movies or any other piece of artwork that can be represented in a planar surface. There are 12 properties defined and stored in the “display set” text files:

<b>ID</b>	Key identifier
<b>Node_ID</b>	Node Key identifier
<b>Width</b>	Represents the physical width of image in pixels.
<b>Height</b>	Represents the physical height of image in pixels.
<b>Real Width</b>	Represents the real world width of the image or painting.
<b>Real Height</b>	Represents the real world height of the image or painting.
<b>Image path</b>	System path where the image is stored
<b>Display description</b>	Description to be used in the scene graph next to the display. For example, “Las Meninas” or “The Family of Philip IV”.
<b>Relative X rotation</b>	X-rotation to apply around the x axis in the middle point relative to the image, in the scene graph.
<b>Relative Y rotation</b>	Y-rotation to apply around the x axis in the middle point relative to the image, in the scene graph.
<b>Relative Z rotation</b>	Z-rotation to apply around the x axis in the middle point relative to the image, in the scene graph.
<b>Relative X</b>	Translation of the display over the x-axis.



<b>position</b>	
<b>Relative Y position</b>	Translation of the display over the y-axis.
<b>Relative Z position</b>	Translation of the display over the z-axis

#### 4.1.2 Panoramic Displays

Panoramas can be represented by a 360°x180° spherical picture panorama or a cylindrical or any other non-planar display. For this work we used the POV-Ray[POV] tool and we made scripts to produce images of spherical panoramas. We then mapped those images in 3D spherical models specially modelled for this kind of nodes. We also used spherical panorama images available under free public licenses, or under author consent.

#### 4.2 Node Structure

A node is a piece of 3D Geometry in which the displays will be placed. It can be a 3D object modelled in Blender [BL] or any other 3D application, that will host the displays in a higher level node in the scene graph, for example a room space. Nodes are defined by 10 properties for each node inside the text file:

<b>ID</b>	Key identifier
<b>Name</b>	Name of the node. For example, “Velazquez 1620 - 1630”. Can be used as text in the scene graph.
<b>Description</b>	Description of the node. Can be used as text in the scene graph.
<b>Category</b>	Category of the node. Used to classify different classes of nodes.
<b>3D Geometry file name</b>	Path + file name of physical disk location of the model file. For example “images\room4.3ds”.
<b>X rotation</b>	X-rotation to apply around the x axis in the central point relative to the object in the scene graph.
<b>Y rotation</b>	Y-rotation to apply around the x axis in the central point relative to the object in the scene graph.
<b>Z rotation</b>	Z-rotation to apply around the x axis in the central point relative to the object in the scene graph.
<b>X position</b>	Translation of the object over the x-axis.
<b>Y position</b>	Translation of the object over the y-axis.
<b>Z position</b>	Translation of the object over the z-axis.

### 5. CONTENT DEVELOPMENT AND TESTING

#### 5.1.1 Content Development

A number of tools were used for developing content in the Crystal Ball framework:

Microsoft® Visual Studio .Net 2003™ with C++ was used to program the base scene graph. The programming was done in Visual C++ with Openscenegraph libraries, generating a gallery structure and an OSG file.

Autodesk® 3ds Max 9™ [3DS] was used to author the 3D models of the nodes and links (see Figure 11 and Figure 17). This tool was also used to perform some 3D modelling conversions.

The Blender 3D modelling tool was used to give more detail to the 3D models of nodes and links.

OpenSceneGraph 2.0 [OSG], was the selected object-oriented framework on top of OpenGL, that manages the scene graph of the system. It includes an API to create and transform geometry elements, animations, visual effects and supports a set of 3D functions and plug-ins.

GtkRadiant 1.5.0 [GTK] (Fig. 14), is the adopted tool to create or edit ready-to use maps for the CaveH environment. This is a well known tool of Quake III game community and was adapted for the purposes of the CaveH platform. It supports reading of 3D file formats produced in several modelling applications. GtkRadiant also supports the definition of triggered actions.

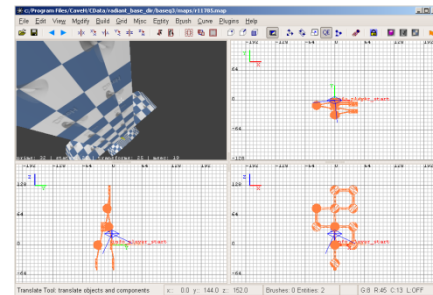


Figure 14 – GTK Radiant screenshot

OSG objects to be placed in the 3D scene can be generated in two different ways:

- Through C++ programming, using classes available in the OSG libraries, exporting the scene graph to a .osg file, containing geometry, lights, smoke or other visual effects.
- By modelling in Autodesk® 3ds Max 9™ or Blender and exporting the result model to the OSG format.

CaveHSpawner (Fig. 15) is an interface for a series of tools to create, edit and run ready-to use maps in the CAVEHollowspace. CaveHSpawner will compile the map produced by GTK Radiant into an octree structure in OSG format, covering all referenced objects and properties and producing as the output, a game environment like Quake 3. This environment will be distributed to the slave nodes of the multi-projection system of the CAVEHollowspace, in the topology described in the preceding paragraphs, using the CaveH middleware.

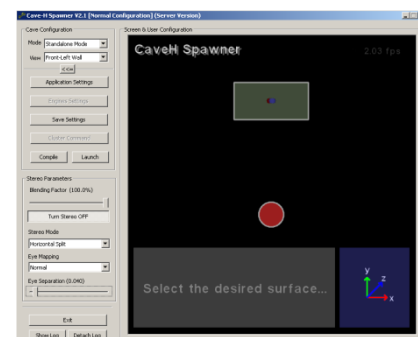


Figure 15 – CaveHSpawner screenshot

POVRay[POV], is an open source ray tracing tool. It has a powerful scripting language that uses complex math functions to create complete ray traced 3D scenes. It was used to generate 3D panoramas used on spherical gallery

panoramas. A 360°x180° panoramic virtual camera was used to make an image mapped on the spherical display object (Fig. 16).

Gimp[Gimp] is an Open Source tool used for picture/photo effects and adjustments. We have used this tool to improve textures and image colours.

The MeshLab[Mesh] tool was used to make 3D model corrections, adjustments and format conversions.

Last but not least, the Second Life® platform with LSL (Linden™ Scripting Language), was used to make a preliminary study and an essay on the plan of operation of virtual galleries. It helped to compose the storyboard of the Crystal Ball immersive experience.

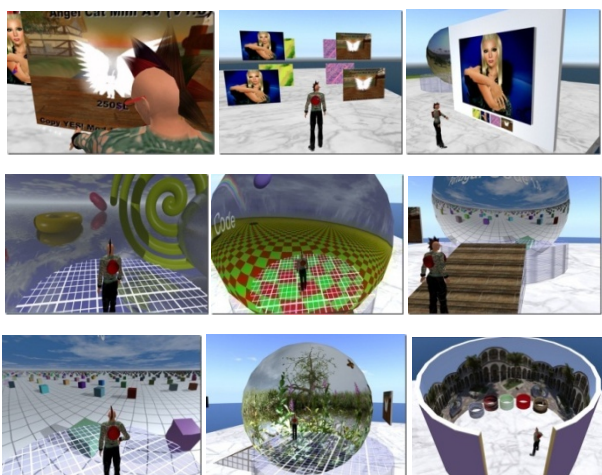


Figure 16 – Storyboard. [SL09] Second Life® World Pictures – Rui de Almeida.

Second Life® World Modelling and Scripting – Rui de Almeida.

Image panoramas(screens 4,5,6) “PT-Code1” and “PT-Code2” and POV Ray Programming – Rui de Almeida.

Image panorama(screen 9) “Patio” in POV Ray by Jaime Vives Piqueres.

Image panorama(screen 8) “The Dark Side Of Trees” in POV Ray by Gilles Tran.

## 5.2 Preliminary Results

To assess the applicability of CBG, we have performed two experiments: a desktop PC experiment and an immersive Virtual Reality System (CaveH [Dias 07]) one.

### 5.2.1 Desktop PC Experiment

The first navigation experience was made by nine unpaid male subjects, selected for the undergraduate course of ISCTE-IUL, using Microsoft Windows XP® on a desktop PC, equipped with a 22-inch monitor and a 2 button + wheel mouse (Fig. 17). To navigate through the scene the subjects have used the CaveH PC viewer and also the Openscenegraph viewer. The Second Life® viewer was used for some earlier experiments. There was nothing new in what concerns to sense of immersion, as experienced by the subjects. It was a common 3D video game like experience. The mouse and keyboard navigation paradigm always left us with a sensation of being out of the scene, and being only an outside observer.



Figure 17 – CBG Desktop PC screenshots.

We have not detected any performance, colour or image processing problem. In general, image quality was perceived as good, but not wide enough for a Virtual Reality experience. It was a poor interaction experience when compared with live painting observation in a real exhibition.

### 5.2.2 Immersive Virtual Reality Experiment

The CAVE™ [Cruz-Neira93] is a room-sized, high-resolution 3D video and audio environment invented in 1991 at Electronic Visualization Laboratory of University of Illinois. It was created by Carolina Cruz-Neira, Dan Sandin and Tom DeFanti, along with other students and staff. Graphics are projected in stereo onto several planes, such as walls and the floor, and viewed with stereo glasses that can be equipped with a tracking sensor. As the user moves within the display boundaries, the correct perspective is displayed in real-time to achieve a fully immersive experience.

#### 5.2.2.1 CAVE Hollowspace at Lousal [Dias07]

The CAVEHollowspace system, that follows the basic architecture of the CAVE™, is located at Lousal Mine in the Alentejo Region in Portugal that was operated between 1900 and 1988. The physical configuration of the CAVEHollowspace of Lousal, includes six projection planes that make a U-shaped projection surface (Fig. 18): there are two retro-projected planes in the front (5.6x2.7 m), two front-projected planes in the pavement (5.6x2.7 m) and a retro-projected plane on each side (3.4 m 2.7 m).

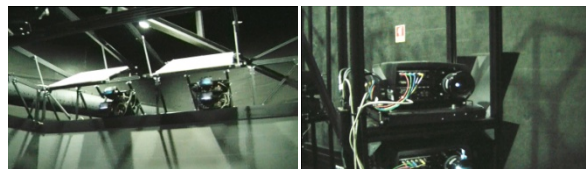


Figure 18 – CaveH projectors.

For stereo viewing, the Infitec passive system [INFITEC] is used (Fig. 19). There are seven speakers that distribute 3D spatial sound over a 7.1 topology, including a 800 watt subwoofer. The system is controlled by a network cluster of 8 computers running Microsoft operating systems and connected by a Gigabit Ethernet.





Figure 19 - CaveH Stereo projection and Infitec passive glasses and CaveH stereo projection.

#### 5.2.2.2 CAVE Hollowspace Experience

The same nine subjects, who had previously experienced the system as a desktop PC, went to Lousal to experience the immersive system. Although this was a new experience for them, unlike the use of desktop PC, all were unanimous in referring an emerged and more involved feeling of immersion in the art gallery visiting experience, even to the point of feeling nausea when navigation was subject to rapid changes in direction and speed.

### 6. CONCLUSIONS

The ultimate immersive 3D Virtual Reality experience would be in an environment where we could walk, run, look around and feel the surrounding sound and other sensorial input, as we do in real life environments. This is definitely not possible in front of a desktop PC screen. For sure this is not yet fully possible in CAVE™ type of environments, but in this case, in addition to the psychological engagement, there is also an increased sensorial and a physical engagement.

In this paper we have presented Crystal Ball, as a concept and a tool suite for authoring, configuring, organizing and displaying 3D virtual galleries, depicting multimedia content, like art, picture, movie or any other kind of media. Our preliminary ad-hoc studies carried with nine students of ISCTE-IUL, gave us some hint on the superiority of large scale immersive virtual environments, such as the CAVEHollowspace (CaveH) of Lousal, in experimenting a visit to a virtual gallery produced by Crystall Ball, if we compare it to a standard PC virtual navigation experience. In fact, in CaveH, we can look and walk around, having the full perception of space and dimensions in 3D. At the same time we can feel the surrounding sound. All of this gives a different sensorial stimulation and a completely different 3D Virtual Reality experience. More complete usability evaluation experiments, involving a panel of test subjects and analysis of completion of task assignments via objective metrics is needed, to further extract conclusions about the benefits brought by immersive environments to the “virtual gallery visit” metaphor. In 3D Virtual Worlds, the user represented by an avatar, although having an immense amount of content, social networking engagement and a myriad of tools and possibilities for communication marketing, business, and entertainment, still is kept as an outsider observer over the world. Future human-computer interaction and visualization paradigms will bring the possibility of making available online virtual environments, more close to CAVE™ type of immersive systems, and the authors believe that Crystal Ball is a contribution in that direction.

### 7. REFERENCES

- [BL] Blender Foundation <http://www.blender.org>
- [Cruz-Neira93] Cruz-Neira, C., Sandin, D., DeFanti, T. “Virtual Reality: The Design and Implementation of the CAVE®” *Proceedings of SIGGRAPH 93 Computer Graphics Conference*, 08/01/1993 - 08/01/1993
- [Dias07] Dias, M. S. et al., “CaveHollowspace do Lousal – Princípios Teóricos e Desenvolvimento”, *Curso Curto, 15 Encontro Português de Computação Gráfica, Microsoft, Porto Salvo*, Out 2007.
- [Frias09] Frias P., “Palup Ling in SL”, *Communication Sciences Department - Oporto University* <http://discursosdooutomundo.blogspot.com/>
- [Gimp] GIMP - The GNU Image Manipulation Program <http://www.gimp.org>
- [GTK] GtkRadiant, Maio 2008. <http://www.qeradiant.com>
- [INFITEC] Infiscape, “INFITEC Stereo Display Technology”, 2009 <http://www.infitec.net/infite.html>
- [Kaczmariski05] Kaczmariski, H., Goudesuene, C., Schaeffer, B., Chong, L., Marshack, R., Hendrickson, L., Crowell, J. “CANVAS: A Virtual Reality Environment for Museums”, *EVA Firenze 2005*, “Application Framework for CANVAS, The Virtual Reality Environment for Museums”, *EVA London 2005*.
- [Lu08] LU, W., Zeng, D., Pan, J., “An XML-Based Scene Description Language for 3D Virtual Museum”, *Department of Computer Science and Technology, Nanjing University, Nanjing*
- [Marshack07] Marshack, R., Duchnowski, N., Watt, A., Jackson, J., Kaczmariski, H., “Creating Immersive Art Without a Programmer: The first year for CANVAS, A Virtual Reality Environment for Museums”, *EVA Firenze 2007*
- [Mesh] Meshlab, <http://meshlab.sourceforge.net/>
- [OSG] OpenSceneGraph, Jan 2009. <http://www.openscenegraph.org/>
- [POV] POV-Ray The Persistence of Vision Raytracer <http://www.povray.org/>
- [Schendel08] Schendel, J., Liu, C., Chelberg, D., Franklin, T., “Virtual Gallery Walk, an Innovative Outlet for Sharing Student Research Work in K-12Classrooms”, *Ohio University*
- [Semiao08] Semião, P. M., Carmo, M. B., “Virtual Art Gallery Tool”, GRAPP 2008 - International Conference on Computer Graphics Theory and Applications, Faculdade de Ciências da Universidade de Lisboa,
- [SL09] Virtual world Second Life <http://www.secondlife.com>
- [Urban07] Urban, R. et al., “A Second Life for Your Museum: 3D Multi-User Virtual Environments and Museums”, in J. Trant and D. Bearman (eds.). *Museums and the Web 2007: Proceedings, Toronto: Archives & Museum Informatics, published March 1, 2007 Consulted January 26, 2010.* <http://www.archimuse.com/mw2007/papers/urban/urban.html>.
- [3DS] Autodesk - Autodesk 3ds Max, 2009. <http://usa.autodesk.com/>

Aveiro, 13, 14 e 15 de Outubro de 2010

# Aplicação de personagens virtuais como guias de visita

Pedro Ribeiro  
Mestrado em Informática, Univ. do Minho  
Campus Gualtar, 4710-057 Braga, Portugal  
pedroribeir@gmail.com

Tiago Silva  
Mestrado em Informática, Univ. do Minho  
Campus Gualtar, 4710-057 Braga, Portugal  
tmgsilva@gmail.com

Rui José  
Mestrado em Informática, Univ. do Minho  
Campus Gualtar, 4710-057 Braga, Portugal  
rui@dsi.uminho.pt

## Sumário

*Tirar o maior aproveitamento possível da experiência de visita em espaços como museus ou exposições tem sido alvo de muitos estudos. Neste projecto foram exploradas as potencialidades de comunicação de plataformas de personagens virtuais como tecnologia de suporte a visitas. Em particular, investigou-se até que ponto uma personagem virtual pode desempenhar funções próximas daquelas que normalmente se atribuem a um guia Humano, ou seja, ser capaz de acompanhar pequenos grupos, transmitindo em cada ponto da visita as informações adequadas a esse local. Como principal conclusão foi identificado que o cenário em que a utilização deste tipo de sistemas se poderá mostrar mais vantajosa será no caso de visitas de estudo com professores, ou outras em que um dos elementos do grupo, mesmo não conhecendo o museu, tem objectivos de ensino para com os restantes. Nesse caso, a não existência de um guia Humano tem como vantagem potencial permitir ao grupo de visitantes definir o seu ritmo de visita e ao mesmo tempo facilita o contacto social entre os elementos do grupo, que é um factor bastante importante para a aprendizagem. O sistema aproveita as potencialidades de comunicação das personagens virtuais para transmitir a informação, mas o facto de não exigir um guia Humano a narrar o guião dá alguma liberdade aos vários elementos para terem conversas informais sobre os conteúdos da visita.*

## Palavras-chave

*Personagens virtuais, experiência de visita, storytelling*

## 1. INTRODUÇÃO

A experiência de visita a espaços expositivos, como museus ou galerias, tem sido objecto de muitos estudos ao longo dos últimos anos. Em particular, a utilização de novas tecnologias de informação tem proporcionado muitas oportunidades para a concepção de novos tipos de experiência que promovam um maior desenvolvimento de conhecimentos no contexto da visita. Contudo, a introdução dessas tecnologias no sentido de enriquecer a experiência de visita a diversos níveis, nomeadamente ao nível pedagógico, lúdico e social, e tendo em conta os diferentes tipos de visita que podem existir (temáticas, escolares, grupos familiares, entre outras) representa necessariamente um desafio complexo, multi-disciplinar e multifacetado, para o qual é fundamental dispor de um leque alargado de alternativas que possam servir de forma adequada as muitas potenciais aplicações.

Neste artigo exploramos de forma muito específica as potencialidades de comunicação de plataformas de personagens virtuais como tecnologia de suporte a visitas. Em particular, pretendemos investigar até que ponto uma personagem virtual pode desempenhar funções próximas daquelas que normalmente se atribuem a um guia Humano, ou seja, ser capaz de acompanhar pequenos grupos,

transmitindo em cada ponto da visita as informações adequadas a esse local.

Este trabalho foi realizado com base no pressuposto de que o cenário da visita é constituído por diversos pontos de interpretação em salas diferentes e que é visitado por pequenos grupos. Assumimos também que em cada uma dessas salas existirá pelo menos um ecrã onde as personagens virtuais irão aparecer para dar aos grupos da visita as explicações respectivas. Esta abordagem surgiu de uma opção de base deste trabalho em que se optou por explorar uma solução que não tivesse necessidade de dispositivos pessoais como telemóveis ou outros. Uma opção desse tipo seria interessante por permitir acompanhar mais facilmente os visitantes durante todo o percurso da visita. Contudo, dado o foco nas visitas de grupo, optou-se pelos ecrãs públicos pois permitem uma experiência partilhada entre os visitantes de um mesmo grupo em vez de os alienar na consulta dos seus dispositivos pessoais.

### 1.1 Objectivos

No âmbito do objectivo genérico que foi apresentado, este trabalho tem os seguintes objectivos:

- Desenvolver e avaliar um sistema de apoio a visitas capaz de suportar requisitos genéricos

desse tipo de sistemas e integrar interacção baseada em personagens virtuais;

- Identificar que mais-valias potenciais podem existir na utilização de uma plataforma de personagens virtuais e, portanto, como é que a experiência de visita pode ser melhorada recorrendo a personagens virtuais;
- Identificar os modelos de interacção que melhor se adequem a uma visita baseada em personagens virtuais;
- Estudar o impacto, contributos e limitações que as diversas tecnologias necessárias para um sistema deste género poderão ter na sua realização.

A metodologia seguida consistiu em estudar cenários semelhantes ao cenário alvo do nosso trabalho, identificando assim alguns requisitos base. De seguida, desenvolveu-se um sistema de apoio a visitas baseado em personagens virtuais e implementou-se um protótipo funcional para poder realizar uma primeira avaliação do sistema. Com base nesse protótipo foram realizados testes num contexto de um espaço físico de visita simulado, para aferir sobre a perspectiva dos utilizadores e obtenção de resultados preliminares em relação a aspectos centrais do trabalho como sejam o modelo de interacção, a usabilidade e de uma forma geral a influência do sistema na experiência de visita. Os resultados mostraram uma grande receptividade em relação ao conceito e sugerem que este será um modelo capaz de se adaptar a diversos tipos de visita, espaços físicos e grupos de visita, proporcionando uma experiência de visita personalizada numa perspectiva complementar aos modelos de visita actualmente utilizados. A utilização de personagens virtuais adiciona valor acrescentado pelo facto de permitir uma maior personalização da visita, bem como dar maior liberdade aos grupos de visita para comunicar e realizar visitas ao seu ritmo e com a sequência desejada.

## 2. TRABALHO RELACIONADO

A experiência de visita a museus é uma área muito vasta e com um corpo de conhecimento já bastante desenvolvido. Nesta secção apresentamos apenas alguns sistemas cujas características são especialmente relevantes para os objectivos específicos do nosso trabalho. Uma dessas abordagens é apresentada em [Rocchi04], onde os autores propuseram um sistema para visitas de museus. Este sistema é constituído pelo dispositivo móvel do visitante e por ecrãs públicos de maior dimensão, localizados em pontos específicos do museu. No dispositivo móvel são feitas apresentações personalizadas através de personagens virtuais, que são geradas dinamicamente pelo sistema. Os ecrãs públicos têm como função disponibilizar informação adicional. O input que o sistema tem provém da localização do visitante bem como das análises que o sistema faz aos comportamentos do visitante e informação já apresentada.

A abordagem utilizada nas apresentações é baseada no uso de personagens virtuais. São produzidos automaticamente pequenos vídeos para serem exibidos nos pequenos

ecrãs dos dispositivos móveis, onde os personagens virtuais se comportam como apresentadores. A construção da animação das personagens apoia-se em regras de apresentação de documentários, onde as personagens têm como objectivo ajudar a resolver certos problemas típicos dos visitantes.

No início da visita os visitantes têm a possibilidade de escolher a sua personagem, sendo que estas são usadas como metáforas para distinguir os interesses dos visitantes no contexto da visita. O visitante tem ainda controlo sobre a apresentação, ao ponto de poder interromper ou ignorar uma determinada apresentação. Este sistema possui por fim a capacidade de reconhecer quando o visitante se aproxima de um ecrã, fazendo automaticamente a transição da personagem visível do ecrã do dispositivo móvel, para o ecrã do museu que está próximo do visitante.

Um dos pontos importantes deste projecto é a utilização de personagens virtuais como meio de comunicação. Este aspecto foi abordado no trabalho “*Design of a Virtual Human Presenter*” relatado em [Noma00]. Os autores desenvolveram um sistema com o objectivo de criar um apresentador virtual para contextos genéricos. Uma das temáticas que eles abordam é a forma como uma apresentação deve ser feita ao nível da postura, gestos e contacto visual. Eles tomam a postura como um dos mais preponderantes elementos visuais numa apresentação. Assumem que o apresentador deve orientar sempre os ombros para a audiência (ponto onde a câmara virtual se encontra). O sistema desenvolvido permite também que o apresentador possa orientar-se para que possa facilmente apontar para outros elementos virtuais que este pretenda explicar. A composição visual que os autores expõem coloca o apresentador de corpo inteiro de um dos lados do ecrã e outros elementos que este pretenda apresentar aparecem do outro lado do ecrã.

Na abordagem que estamos a explorar o aparecimento das personagens faz-se apenas em ecrãs públicos, não sendo necessário qualquer tipo de dispositivo pessoal caro. As personagens vão “acompanhando” os visitantes ao longo do percurso, aparecendo nos ecrãs e permitindo dois modos de interacção: interacção implícita, pela entrada na sala, e interacção explícita, fazendo utilização de marcadores visuais.

## 3. SISTEMA DE APOIO A VISITAS

Um dos objectivos deste trabalho foi também estudar de que forma uma plataforma de personagens virtuais poderia ser integrada no contexto de sistemas digitais de apoio a visitantes. Nesse sentido, e também para ter um sistema que servisse de base à avaliação com utilizadores, desenvolvemos o nosso próprio sistema de apoio a visitantes com base em personagens virtuais, que será descrito nesta secção.

### 3.1 Requisitos

Com base na análise de cenários semelhantes ao proposto para o nosso sistema, definimos um conjunto de requisi-

tos que o sistema deveria ser capaz de cumprir, nomeadamente:

- Identificar presença dos grupos e disponibilidade para ouvir o discurso;
- Ser sensível a planos flexíveis (ter em conta que os grupos de visitas podem abandonar os pontos de interacção);
- Perceber quando é que os grupos pretendem parar ou interromper o discurso;
- Ser robusto a diversos percursos/sequências de visita;
- Permitir ser configurado com diversos tipos de conteúdos (guiões de visita);
- Suportar diversos tipos de grupos de visita (diferentes idiomas, por exemplo).

### 3.2 Arquitectura

De uma forma geral o sistema tenta enriquecer a experiência de visita recorrendo à utilização de ecrãs dispostos nas diversas salas de um museu. Nos ecrãs aparecem personagens virtuais que têm como objectivo interagir com os visitantes do museu. Essa interacção é adaptada ao contexto da visita e comportamento dos visitantes, recorrendo a mecanismos de sensores e a informação do estado da visita que vai sendo guardada ao longo de toda a visita.



Figura 1 – Arquitectura lógica do sistema

O módulo de *sensores* comunica com o módulo de **estado da visita** e este último, por sua vez, fornece ao módulo de *storytelling* a informação necessária para que este consiga controlar a sequência de narrativas que cada personagem (módulo de **interacção**) deverá apresentar.

### 3.3 Sensores

O mecanismo de sensores utilizado neste sistema pode ser dividido em duas categorias: sensores implícitos e sensores explícitos.

Os sensores implícitos são utilizados para que o sistema consiga detectar que um grupo de visitantes está perto de um ecrã. Este tipo de sensores utiliza *tags* RFID para identificar os diferentes grupos de visitantes. A *tag* RFID deve estar integrada no objecto que é fornecido no início de visita a cada grupo de visitantes.

Os sensores explícitos são utilizados pelo sistema para perceber se um determinado grupo de visitantes pretende interagir ou não com a respectiva personagem virtual. Este mecanismo de sensores utiliza visão por computador para reconhecer marcadores que estão contidos no objecto que cada grupo de visitantes possui (fornecido no início da visita). O objecto tem dois marcadores e as personagens virtuais apresentam comportamentos adequados para que o grupo de visitas perceba quando os pode utilizar. Um dos marcadores é utilizado para indicar que o grupo pretende iniciar a interacção com a personagem e o outro para indicar que o grupo pretende parar ou interromper a interacção com a personagem. A disposição dos marcadores no objecto é feita de forma a evitar que os grupos de visitantes mostrem os marcadores acidentalmente ou mostrem os dois marcadores em simultâneo. O planeamento realizado no início do projecto apontava para a utilização da plataforma de realidade aumentada ARToolkit para detecção visual de marcadores. No entanto, tendo em conta que o marcador visual do ARToolkit tem uma configuração visual muito específica, decidiu-se procurar uma plataforma alternativa que ultrapassasse esta exigência ao nível do aspecto dos marcadores. Uma plataforma para a interpretação de qualquer imagem como sendo um marcador visual. Optou-se pela utilização do algoritmo SURF (*Speeded Up Robust Features*) recorrendo à biblioteca OpenCV. Através deste foi possível garantir a criação e incorporação de marcadores visuais tendo em conta qualquer linha de design. Estes dois modelos de sensores permitem criar um modelo genérico de interacção combinado, que não depende do tipo de objectos utilizados. O objecto pode ser concebido e adaptado ao espaço físico da visita e estar até de alguma forma relacionado com a própria temática da visita. Um exemplo disso seria um livro, no caso de uma visita a uma biblioteca, ou então uma bola, numa visita a um estádio de futebol.

### 3.4 Storytelling

O módulo de *storytelling* tem como função seleccionar e sequenciar as narrativas de cada personagem.

Tendo em conta a informação recebida a partir do módulo de *sensores*, o estado da visita e o ecrã/sala em que o sistema está a ser utilizado, o módulo de *storytelling* decide qual o segmento de narrativa a apresentar para o grupo de visitas que se encontra actualmente a interagir com o sistema.

As narrativas de *storytelling* são guardadas em ficheiros BML (*behaviour markup language*) que posteriormente são lidas pela plataforma de personagens virtuais utilizada neste projecto.

Cada personagem tem um ficheiro BML associado onde são descritos todos os seus comportamentos. A parte inicial deste ficheiro é reservada à especificação da personagem, como o seu nome, a malha poligonal a usar e as características da sua voz. Os seus comportamentos estão divididos por eventos. Ao evento está associado o seu nome, bem como um campo que possibilita uma das características mais importantes do sistema, que é a defi-

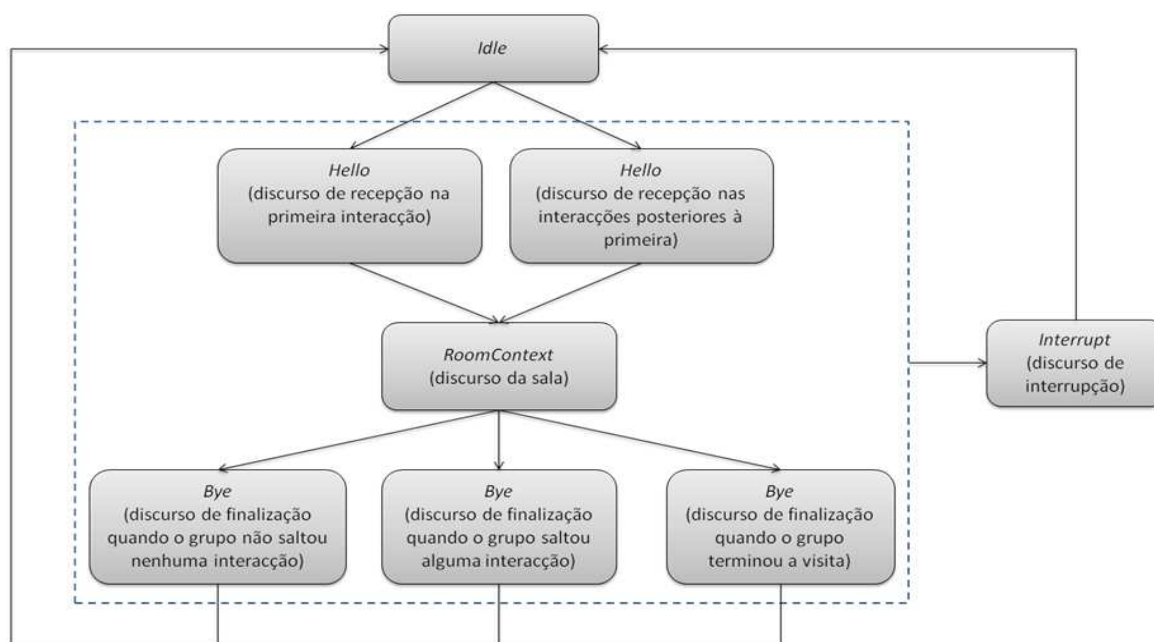


Figura 2 – Máquina de estados do *storytelling*

nição do suporte gráfico para o discurso. A associação do suporte gráfico ao evento possibilita ao sistema sincronizar o discurso de uma personagem com o aparecimento de imagens associadas, um pouco à semelhança do que acontece com o *pivot* de um noticiário.

Os diversos tipos de eventos contidos nos ficheiros BML estão directamente relacionados com a máquina de estados *storytelling* (ver figura 2). Os tipos eventos contidos nos ficheiros BML são:

1. Evento **Idle** – Relativo ao estado de espera da personagem, quando este aguarda uma interação explícita do grupo de visitantes. Aqui a personagem tem uma atitude passiva, sorrindo sempre que olha na direcção dos visitantes e demonstrando vontade de interagir com o grupo;
2. Evento **Hello** – Este evento é despoletado quando o grupo decide interagir explicitamente (utilizando marcadores visuais). Neste evento, a primeira acção que a personagem concretiza é a mudança de posição para uma “mais próxima dos visitantes”. A personagem passa a ocupar uma porção maior do ecrã. Logo de seguida a personagem saúda os visitantes de uma das duas maneiras possíveis: se for a primeira interação que tem com o grupo de visitas, então dá as boas vindas tendo em conta que o grupo está a interagir com o sistema pela primeira vez, referindo o nome do grupo (personalização do discurso) e introduzindo de seguida a sala em que o grupo se encontra, caso contrário, apenas cumprimenta o grupo, dando as boas vindas pelo facto de estar a interagir na sala em que se encontra naquele momento;
3. Evento **RoomContext** – Este evento é despoletado logo depois do evento **Hello**. Aqui a personagem tem

um discurso que corresponde à interpretação da sala em que o grupo de visitas se encontra a interagir naquele momento. Este discurso corresponde directamente ao conteúdo do guião da visita;

4. Evento **Interrupt** – Sempre que os visitantes mostrem à personagem o marcador visual de interrupção, o evento **Interrupt** é disparado. A primeira acção que a personagem realiza é a mudança de posição para uma “mais afastada dos visitantes”. Simultaneamente a personagem tem um comportamento de espanto e profere algumas palavras de resignação. Logo depois a personagem retoma o evento de **Idle**.
5. Evento **Bye** – O evento **Bye** é lançado sempre após a conclusão do discurso de uma determinada sala (evento **RoomContext**), se não houver qualquer tipo de interrupção. Este pode ser feito de três formas, tendo em conta o estado da visita:
  - a. Se os visitantes não alteraram a sequência normal da visita, então a personagem despede-se apenas indicando que espera “ver o grupo” na próxima sala;
  - b. Se os visitantes não interagiram com o sistema numa determinada sala e, portanto, não seguiram a sequência normal da visita, então a personagem despede-se também indicando que espera “ver o grupo” novamente, mas avisa que possivelmente não visitaram uma ou mais salas. As salas que o sistema considera não terem sido visitadas correspondem àquelas que o grupo não interagiu com o sistema;
  - c. Se os visitantes estão na última sala e já interagiram e visitaram todas as salas, então a personagem despede-se definitivamente do grupo.



A cada um destes eventos estão também associados uma série de comportamentos como a narrativa, direcção do olhar, bem como emoções faciais das personagens virtuais.

Na Figura 3 pode-se ver um excerto de um ficheiro BML que corresponde ao evento **RoomContext**.

```
<evt event="RoomContext" subject="UbiVisit">
  <speech id="s1" actor_id="Dave" text="Hi guys!
welcome to the Open Space room!"/>
  <facegaze id="fg16" absolutetime="500"
duration="2000" actor_id="Dave" target="visitors"
angle="" direction="" type="face" amount="0.2"
shape="HAPPY_SHAPE" />
</evt>
```

Figura 3 – Excerto de BML

### 3.5 Plataforma Creactor

Um dos pressupostos deste projecto como já foi referido anteriormente é a utilização de personagens virtuais. Como tal, é necessária uma plataforma que se encarregue dessa tarefa. No âmbito do projecto VirtualActor [Turge109], está sendo desenvolvido neste momento um software de nome Creactor, cujo propósito é a criação de actores virtuais. O utilizador transforma-se portanto num realizador de um filme de animação. O Creactor é um software genérico, que inclui editores e um sistema de animação para personagens virtuais. Os editores permitem o uso de diferentes técnicas de controlo e de IA Inteligência Artificial) para a concepção e edição dos actores virtuais com os seus comportamentos. O software gera animações em 3D com actores virtuais que são emocionalmente expressivos, com aptidão para agir com autonomia e capazes de se expressarem emocionalmente, com voz e com sinais não-verbais. O Creactor recorre à plataforma de *text to speech* Loquendo, para realizar a síntese da voz dos actores.

### 3.6 Ocupação do ecrã

A forma como os elementos visuais são organizados no ecrã baseia-se na metáfora do apresentador.

Os elementos visuais que a organizar são os actores associados aos grupos que se encontram na zona de interacção, bem como o material de apoio que o actor utilizara para enriquecer a apresentação, como imagens de artefactos presentes naquela sala do museu.

O material de apoio a apresentação é visível no centro do ecrã. Já os actores estão posicionados no primeiro ou terceiro terço da imagem.

Esta disposição de elementos visuais no ecrã permite ao sistema apresentar até dois actores virtuais no ecrã, para os casos em que o mesmo spot de interacção está ocupado por mais do que um grupo de visitantes.

O apresentador (actor) activo estará sempre mais próximo dos visitantes (maior dimensão), e será o único que falará.



Figura 4 – Screenshot da aplicação

### 3.7 Múltiplos grupos de visita

Durante a interacção de um grupo de visitantes com o sistema, pode-se aproximar mais um grupo do ecrã que entretanto já está ocupado. Neste caso a personagem desse grupo surge no ecrã de forma a mostrar que o sistema também tem interesse em interagir com ele. Este comportamento não influencia o discurso verbal do actor que já estava em cena. Se existirem mais do que dois grupos num ecrã, o que será pouco provável devido ao facto de existirem limitações de espaço físico em frente ao mesmo, o sistema ignora os últimos grupos a aproximarem-se do ecrã.

## 4. AVALIAÇÃO

Nesta secção iremos descrever a primeira avaliação do sistema realizada com base num protótipo funcional.

### 4.1 Instalação do protótipo

Na figura 5 é apresentado o diagrama que mostra a arquitectura física do protótipo que foi instalado para efeitos de avaliação. A arquitectura é baseada no modelo cliente-servidor.

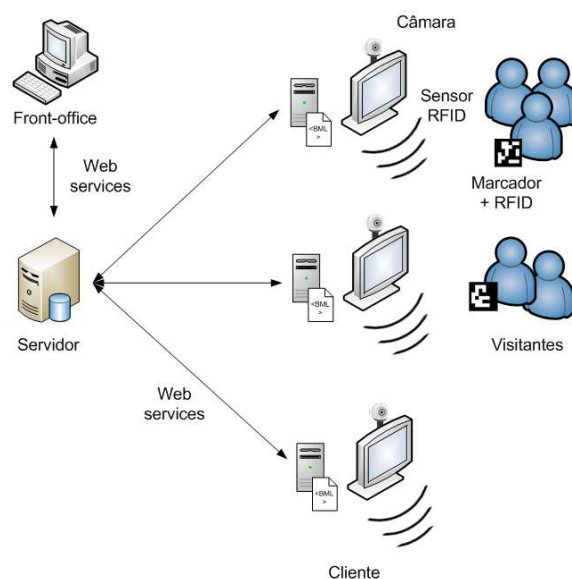


Figura 5 – Arquitectura física do sistema



Esta arquitectura foi implementada no protótipo deste projecto, embora algumas componentes fossem simuladas, mais concretamente os sensores implícitos do sistema, que foram simulados manualmente.

O módulo de servidor é responsável por disponibilizar os *Web Services* e contém também a base de dados onde vão ser guardados os estados das visitas de cada um dos grupos que num determinado momento estão a visitar o espaço físico.

Cada sala tem um módulo de cliente, constituído por um computador, um ecrã, uma webcam e um sensor de RFID. Este módulo é responsável por toda a parte de sensores, personagens virtuais e *storytelling*. Cada módulo de cliente tem localmente os ficheiros BML (*Behavior Markup Language*) necessários para que a plataforma de personagens possa apresentar os diversos estados que as personagens podem tomar, incluindo o discurso específico da sala onde o cliente se encontra.

O módulo de cliente apenas comunica via *Web Services* com o módulo de servidor para obter informação relativa aos grupos de visitantes e ao estado das visitas.

O módulo de *front-office* é utilizado no início da visita para registar toda a informação relativa aos grupos de visita (nome do grupo, idioma da visita, personagem, entre outros). Este módulo comunica também com o servidor utilizando *Web Services*.

## 4.2 Resultados preliminares

Para a avaliação do sistema desenvolvido não foi possível realizar formalmente testes com os utilizadores, no entanto, conseguiu-se expor os utilizadores a uma configuração primária do sistema, com o objectivo de se obter o primeiro feedback participativo.

Esta primeira avaliação do protótipo funcional do sistema decorreu nas instalações do Centro de Computação Gráfica (CCG), que se encontra sediado no campus de Azurém da Universidade do Minho. Realizaram-se diversas experiências de visita, com grupos de voluntários em 3 salas do edifício. Foram necessários quatro computadores, um deles foi utilizado tanto como servidor como spot de visita, enquanto outro serviu única e exclusivamente para efectuar os registos dos grupos de visitantes. Os três computadores a utilizar como spots de visita foram equipados com monitores de 19", colunas de som e uma webcam.

Foram observados 7 visitantes, do sexo masculino com idades compreendidas entre os 23 e os 29 anos. Estes foram distribuídos em 3 grupos de pequena dimensão (entre 2 e 3 visitantes). Os visitantes que se voluntariaram para fazer os testes são informáticos, conheciam o espaço físico da visita, mas não tinham conhecimentos sobre os conteúdos que foram expostos durante a mesma. No início da visita os visitantes dirigiram-se à recepção onde se registou a informação relativa ao grupo, o nome e personagem virtual que acompanhou o grupo durante toda a sua visita. Foi ainda entregue aos grupos no momento do registo, dois marcadores (duas folhas em formato A4) que permitiram aos visitantes interagir explicitamente com as

personagens virtuais. Após o registo, os visitantes interagiram com o sistema em todas as salas, percorrendo as salas do edifício por um trajecto definido e indicado por uma pessoa que os acompanhava e observava. No final da visita, os visitantes foram sujeitos a entrevistas, de forma a ser possível obter resultados para posterior análise, quer a nível da experiência de visita, quer a nível de interacção e usabilidade.

## 4.3 Análise de resultados

No que concerne à interacção e usabilidade, procurou-se obter opiniões relativamente a questão da adequação do tamanho e disposição dos ecrãs bem como a intensidade do som, tendo em conta o grupo e o espaço onde estavam inseridos. Todos os visitantes consideraram que a configuração visual e sonora do sistema se adequava perfeitamente às exigências.

Os visitantes foram consultados no sentido de obter impressões sobre interacção explícita. Genericamente quando lhes foi perguntado se sentiram controlo sobre a aplicação houve um elemento dos sete que respondeu negativamente, alegando que a detecção de imagens não funcionava muito bem. Procurou-se também perceber se os visitantes consideraram o modelo de interacção adequado. Aqui todos responderam positivamente com alguns elogios à flexibilidade e versatilidade desta solução. Foram ainda sugeridas melhorias, quer a nível da velocidade de reacção do sistema aquando da detecção da imagem, quer a nível da necessidade de substituir a folha A4 na qual se encontrava a imagem, por um objecto mais intuitivo (relacionado com o local da visita) e que tivesse outro tipo de manuseamento e robustez.

Um outro foco de estudo está evidentemente na interacção implícita. Nesse sentido averiguou-se se os visitantes sentiram resposta do sistema quando chegaram próximo dos ecrãs. Esta pergunta dividiu os entrevistados: uns sentiram resposta por parte do sistema, outros não se aperceberam de nada, dizendo que as personagens já lá se encontravam quando eles lá chegaram.

Na avaliação da experiência de visita, uma das questões que ficou mais patente através da observação resulta do facto de termos utilizado uma pessoa para conduzir os visitantes ao longo do percurso da visita. Esta decisão acabou por constranger a experiência de visita, quer na tomada de decisões por parte do grupo, na sua autonomia, quer na sequência de visita das salas, que poderia variar de visita para visita.

Tal como é referido por Bitgood [Bitgood02], um dos factores mais importantes para a experiência de visita tem a ver com o contacto social. Bitgood considera este factor muito importante para a aprendizagem informal e possivelmente o aspecto mais importante para a experiência de visita. Nesse sentido decidiu-se perguntar aos visitantes se interagiram com outros membros do grupo ao longo da visita. Apenas um considerou que não interagiu muito com os restantes membros do grupo. Foi ainda observado num dos grupos que, pelo facto de um dos elementos ter conhecimentos aprofundados acerca de um dos projectos

apresentados, fez com que os dois elementos ficassem a conversar sobre o assunto que lhes tinha sido exposto, mesmo depois de terem saído do *spot*.

Um outro aspecto que foi explorado foi a questão da personalização da visita. Os voluntários foram interrogados no sentido de se apurar se eles tinham achado a visita personalizada. As reacções foram bastantes semelhantes, todos consideraram que pelo facto do sistema incluir no discurso da personagem o nome do seu grupo, o sistema de certa forma tornava a visita personalizada. No entanto, foi perceptível pelas suas reacções que esta personalização fica um pouco aquém do desejável.

De notar também que todos os visitantes consideraram o sistema um meio complementar e enriquecedor da visita. Sugeriram no entanto a inclusão de mais conteúdos por cada *spot*.

Por fim, foi colocada uma das principais questões relativas à experiência de visita relacionada com a aceitação que as personagens virtuais têm quando utilizados como guias de visita. Encontraram-se neste campo respostas positivas. Os sete elementos consideraram que as personagens virtuais neste plano têm um enorme potencial. Acharam que é uma forma de comunicação bastante natural e intuitiva. Conseguiram também perceber que as personagens são passíveis de serem adaptados ao contexto da visita e ao próprio grupo.

## 5. CONCLUSÕES

A principal conclusão a tirar deste projecto é o facto da utilização de personagens virtuais ter sido bem aceite pelos visitantes e ter sido reconhecida como uma mais-valia na experiência de visita, na medida em que tornou a experiência de visita mais lúdica sem pôr em causa a transmissão de conhecimento sobre o local da visita.

No que se refere ao sistema proposto, este mostrou boa capacidade de adaptação a diversos tipos de conteúdos e mostrou funcionar bastante bem em grupos de pequena dimensão.

Relativamente ao discurso das personagens, pode-se concluir que o conteúdo das narrativas é muito importante para captar a atenção dos visitantes, no sentido de compensar por vezes a falta de expressividade, emoção e fraca linguagem corporal das personagens virtuais (a plataforma de personagens virtuais utilizada apenas mostra a cabeça da personagem e não o corpo inteiro). A máquina de *storytelling* implementada revelou-se bastante flexível e adaptada a diversos tipos de conteúdos, no entanto, a criação de ficheiros BML com bons conteúdos (narrativas e emoções associadas) é de grande relevância para o sucesso deste modelo de visita. Por vezes, os discursos não foram muito claros e naturais, possivelmente devido ao facto de serem demasiados rápidos e com uma entoação não muito perfeita, o que significa que a escolha da plataforma de *text to speech* também é muito importante, assim como a sua configuração.

Ao nível da interacção, conclui-se que o módulo de visão por computador implementado neste sistema para reconhecimento de imagens não funcionou muito bem (senso-

res explícitos), pelo facto da velocidade de reacção do sistema ser lenta, exigir bastante capacidade de processamento computacional e condições de luz controladas. Um outro ponto que também se conclui relativamente à interacção é que marcadores/imagens diferentes obtêm respostas diferentes por parte do sistema relativamente à velocidade de reconhecimento. A solução para este problema deverá passar pela optimização da eficiência do módulo de detecção de imagens, mesmo que isto implique a utilização de marcadores visuais com configurações bastante específicas, que possivelmente não tenham qualquer significado por si só para os utilizadores do sistema.

A avaliação permitiu também concluir que o objecto de interacção pode influenciar bastante na experiência de visita e na captação de interesse de utilização do sistema por parte dos visitantes. O objecto de interacção deverá ser algo fácil de transportar, apelativo, estar de alguma forma relacionado com o local da visita, evitar que os grupos de visita consigam mostrar mais do que um marcador/imagem simultaneamente às personagens e ter uma robustez e custo adequados ao desgaste que se prevê pela utilização repetida do mesmo. Conclui-se também que este modelo de interacção combinado (utilização de sensores implícitos e explícitos simultaneamente), juntamente com a utilização de um objecto que pode ser adaptado ao contexto da visita, torna o sistema muito versátil. Inclusivamente foi sugerido na avaliação do protótipo funcional que as imagens utilizadas para interagir poderiam ser estampadas em t-shirts vestidas por um elemento de cada grupo de visitantes.

O tamanho dos ecrãs e condições sonoras proporcionadas durante o processo de avaliação mostrou ser bastante adaptado ao espaço físico utilizado e a grupos de visita de pequena dimensão (tipicamente 3 elementos). Eventualmente a altura dos ecrãs deveria ser melhor adaptada para que todos os elementos do grupo consigam ver a imagem nas mesmas condições, sem qualquer tipo de ocultação por parte dos elementos que se encontram mais próximos do ecrã.

Embora neste protótipo tenha sido simulado, a utilização de RFID parece ser uma boa solução para o sistema ser capaz de detectar, de uma forma implícita, a presença do grupo junto a um *spot* de interacção. Contudo, em alguns casos o sistema deu aos grupos de visitantes a sensação de que o aparecimento das personagens nos ecrãs era despoletado pelo facto de apresentarem o marcador junto ao *spot* de interacção. Isto revela que de alguma forma as personagens virtuais não apresentaram comportamentos que distinguíssem claramente o estado em que apenas têm intenção de mostrar que detectaram a presença do grupo de visita junto ao *spot*, do estado em que estão realmente interessados em interagir com o grupo e que devem fazer com que o grupo perceba que tem de mostrar o marcador/imagem correcto representado no seu objecto de interacção.

Um efeito secundário e imprevisto do modelo de interacção foi que para os utilizadores o guia não os estava propriamente a acompanhar mas sim em todo o lado. Dado

que o guia aparecia num ecrã de uma sala quando as pessoas chegavam a essa sala, estas acabavam por quase nunca ver o ecrã sem que o guia lá estivesse, tendo assim uma sensação de omnipresença relativamente à personagem. Como trabalho futuro iremos estudar de que forma a sensação de acompanhar os visitantes pode ser melhorada.

Finalmente, e como principal conclusão relativamente ao nosso objectivo central que era explorar as potencialidades das personagens virtuais como guias, podemos identificar desde já que o cenário em que essa utilização se poderá mostrar mais vantajosa será no caso de visitas de estudo com professores, ou outras em que um dos elementos do grupo, mesmo não conhecendo o museu, tem objectivos de ensino para com os restantes. Nesse caso, a não existência de um guia Humano tem como vantagem potencial permitir ao grupo de visitantes definir o seu ritmo de visita e ao mesmo tempo facilita o contacto social entre os elementos do grupo, que é um factor bastante importante para a aprendizagem [Bitgood02]. O sistema aproveita as potencialidades de comunicação das personagens virtuais para transmitir a informação, mas o facto de não exigir um guia Humano a narrar o guião dá alguma liberdade aos vários elementos para terem conversas informais sobre os conteúdos da visita.

As circunstâncias do cenário de avaliação não permitiram testar a utilização do sistema por mais do que um grupo de visita simultaneamente, pelo que isso será uma continuação importante para este trabalho. De uma forma geral, uma avaliação num cenário de visita real, menos

constrangido e mais imprevisível, seria um teste importante à flexibilidade do sistema e à sua capacidade de adaptação a contextos de utilização diversos.

## 6. AGRADECIMENTOS

O trabalho descrito neste artigo foi parcialmente financiado pela Fundação para a Ciência e Tecnologia (FCT) no âmbito do projecto VirtualActor, PTDC/EIA/69236 e financiado pela Comissão Europeia, Seventh (FP7 - 2007-2013) no âmbito do projecto GUIDE, 248893.

## 7. REFERÊNCIAS

- [Rocchi04] C. Rocchi, O. Stock, e M. Zancanaro, "The Museum Visit: Generating Seamless Personalized Presentations on Multiple Devices," IN PROCEEDINGS OF INTELLIGENT USER INTERFACES IUI 2004, 2004, pp. 316--318.
- [Noma00] T. Noma, L. Zhao, e N.I. Badler, "Design of a Virtual Human Presenter," IEEE Computer Graphics and Applications, vol. 20, 2000, pp. 79-85.
- [Iurgel09] I.A. Iurgel, R.E. Silva, P.R. Ribeiro, A.B. Soares, e M.F. Santos, "CREACTOR – An Authoring Framework for Virtual Actors," Proceedings of the 9th International Conference on Intelligent Virtual Agents, Amsterdam, The Netherlands: Springer-Verlag, 2009, pp. 562-563.
- [Bitgood02] Bitgood Stephen: Environmental Psychology in Museums, Zoos, and Other Exhibition Centers. In: R. Bechtel & A. Churchman (eds.), Handbook of Environmental Psychology, John Wiley & Sons, 461-480 (2002).

# Avaliação da interacção com Envolvimentos Virtuais através de matrizes de exploração do espaço

Emília Duarte  
UNIDCOM/IADE – Escola Superior  
de Design  
Av. D. Carlos I, 4, 1200-649 Lisboa  
emilia.duarte@iade.pt

Júlia Teles  
Unidade de Métodos Matemáticos  
Faculdade de Motricidade Humana  
Estrada da Costa, 1499-002 Cruz  
Quebrada, Dafundo  
jteles@fmh.utl.pt

Luís Teixeira      Francisco Rebelo  
Laboratório de Ergonomia  
Faculdade de Motricidade Humana  
Estrada da Costa, 1499-002 Cruz Quebrada, Dafundo  
lmteixeira@fmh.utl.pt; frebelo@fmh.utl.pt

## Sumário

*Este estudo teve como objectivo principal a avaliação da interacção com Envolvimentos Virtuais (EVs) através de matrizes de exploração do espaço. As matrizes são o resultado da exploração de um EV, com dois níveis de poluição visual, onde foram aplicados dois tipos de avisos de segurança, estáticos e dinâmicos, tendo ainda sido considerada uma condição de controlo. Participaram neste estudo 150 indivíduos, 30 em cada uma de 5 condições experimentais, aos quais foi pedido que efectuassem uma tarefa, relacionada com avisos de segurança, que implicava a procura de botões distribuídos por diversas zonas do EV. Partindo do pressuposto que os avisos mais salientes (dinâmicos) seriam detectados mais facilmente, foi colocada a hipótese de que existiriam diferenças na exploração do espaço e que essas diferenças estariam reflectidas nas matrizes. O sistema de Realidade Virtual (RV) imersiva – ErgoVR – recolheu automaticamente diversos dados da interacção, entre os quais se destacam a posição do participante no EV, os tempos e as distâncias percorridas. Esses dados permitiram gerar matrizes de exploração do espaço, por condição experimental, contendo indicação das quadriculas pisadas e a frequência com que o foram. O “acordo” da exploração espacial, nas 5 condições experimentais, foi avaliado através da aplicação do Coeficiente de Correlação de Concordância sobre as frequências com que as quadriculas foram pisadas. Os resultados obtidos revelam valores bastante elevados de “acordo” entre as matrizes, o que permite afirmar que a exploração do espaço no EV não foi fortemente influenciada pela condição experimental.*

## Palavras-chave

*Realidade Virtual, interacção, exploração do espaço, avisos de segurança*

## 1. INTRODUÇÃO

Na maioria das situações do nosso quotidiano somos confrontados com inúmeros estímulos (ex. cartazes, sinais, luzes, sons, pessoas), presentes no envolvimento, que competem em simultâneo pela nossa atenção. Porém, apesar de sermos capazes de dividir a atenção por mais do que um estímulo, tanto a capacidade de dar atenção, como os recursos cognitivos usados no processamento activo da informação são limitados [Anderson00]. Nesse sentido, somos selectivos e, por vezes, focamo-nos na procura activa de um certo tipo de estímulo, ignorando os restantes.

Haverá situações em que não dar ou não manter a atenção, numa determinada informação, poderá não ser grave, mas o mesmo não se pode dizer tratando-se de

informação de segurança (ex. avisos e sinais). O processamento de avisos pode ser sequenciado, de forma geral, em três grandes etapas: dar atenção, compreender e agir em conformidade. Se a eficácia na última etapa é considerada, por diversos autores, como a avaliação mais importante [e.g. Kalsher06; Silver99; Wogalter87], aquilo que acontece na etapa da atenção é crucial, no sentido em que pode estancar, logo no início, todo o processo subsequente. Dificilmente poderemos agir em conformidade com um aviso que não detectámos. Adicionalmente, diversos estudos revelaram a existência de uma relação entre a saliência dos avisos e o sucesso na promoção do comportamento consonante [e.g. Wogalter93]. No entanto, em parte, a importância do grau de saliência também está dependente das experiências e conhe-

cimentos prévios dos indivíduos. Por exemplo, a exposição repetida ao mesmo aviso leva à formação de uma memória que, por sua vez, leva a que o mesmo se torne menos saliente para o receptor em resultado de uma habituação [Wogalter06]. É claro que o processo da atenção não é tão simples quanto o aqui exposto, mas estas noções básicas servem para enquadrar o trabalho desenvolvido.

Em termos genéricos, a tendência será dar atenção (ex. olhar, escutar) aos estímulos mais conspícuos, isto é, àqueles que, devido às suas propriedades ou características, se destacam ou são mais proeminentes no envolvimento. A este respeito, estudos recentes sugerem que avisos dinâmicos, por vezes multimodais, habitualmente de base tecnológica, são mais eficientes do que os equivalentes tradicionais (estáticos), uma vez que possuem características que os tornam mais salientes e mais resistentes à habituação [e.g. Wogalter02; Mayhorn03; Smith-Jackson04; Wogalter05]. O resultado pode ser a obtenção de níveis mais elevados de consonância comportamental. A atenção, tal como acontece com outras etapas do processamento da informação, não pode ser avaliada directamente. Assim, a sua avaliação é feita através de variáveis que, potencialmente, reflectem o fenómeno (definições operativas) como, por exemplo, o tempo de detecção/resposta, o comportamento de observação, fixações oculares, diversas avaliações subjectivas (ex. classificar a saliência de diversas alternativas), entre outras.

A Realidade Virtual (RV) tem sido usada, como ferramenta de investigação na área dos avisos de segurança, nomeadamente, na avaliação da influência dessa informação durante evacuações de emergência [Glover97; Ren08; Gamberini03; Tang09]. As suas potencialidades, para esta área de investigação, são diversas, destacando-se o potencial para disponibilizar contextos de uso realistas, bastante interactivos e com elevados níveis de imersividade, podendo beneficiar a validade externa ou ecológica dos estudos. A RV pode também, com custos relativamente baixos, viabilizar o acesso a todo o tipo de contextos, especialmente aqueles de difícil alcance. Com a RV os estudos podem ser repetidos e replicados, de forma sistemática e rigorosa, sem que as variáveis independentes sofram alterações, beneficiando a validade interna dos estudos. Adicionalmente, a RV pode permitir superar alguns dos constrangimentos relacionados com as questões éticas e de segurança dos participantes, que limitam muitos destes estudos [Duarte, *in print*].

Neste contexto, este estudo teve como objectivo principal a avaliação da interacção com Envolvimentos Virtuais (EVs), através de matrizes de exploração do espaço. As matrizes analisadas foram elaboradas com os dados recolhidos automaticamente, durante a simulação em RV, pelo sistema *ErgoVR*, entre os quais se destacam a posição do participante dentro do EV, os tempos e as distâncias percorridas. Foram avaliadas cinco matrizes, resultantes da interacção com EVs com dois níveis de poluição visual, onde foram aplicados dois tipos de

avisos de segurança (estáticos e dinâmicos), tendo ainda sido incluída uma condição de controlo (neutra). Partindo do pressuposto que os avisos mais salientes (dinâmicos) seriam detectados mais facilmente, foi colocada a hipótese de que essas diferenças estariam reflectidas na exploração espacial do EV. Esperava-se que nas condições experimentais onde os avisos são menos salientes (estáticos) e o ambiente se apresenta mais poluído, houvesse uma maior deambulação do participante no EV, originando uma dispersão quanto às quadriculas pisadas (indicador de maior procura, por áreas mais vastas), bem como frequências mais elevadas de “pisadelas” em algumas quadriculas (indicador de hesitação e/ou de necessidade de uma busca mais detalhada naquela zona).

O “acordo” nas matrizes de exploração espacial, para as zonas de interesse, foi avaliado através da aplicação do Coeficiente de Correlação de Concordância (CCC) [Lin89; Barnhart02; Lin02], considerando como variável de interesse a frequência com que as quadriculas foram pisadas. O CCC é um índice de reprodutibilidade, proposto por Lin [Lin89], que permite quantificar o “acordo” entre duas medições efectuadas numa escala quantitativa contínua, podendo também ser usado para medições numa escala quantitativa discreta [Carrasco05]. Este coeficiente tem duas componentes: precisão, que avalia o quanto as observações se afastam da recta de regressão linear ajustada, e a exactidão, que tem em conta o afastamento entre a recta de regressão e a recta de concordância. Para quantificar o “acordo” entre mais de duas medições, diversas generalizações têm sido propostas, nomeadamente o Coeficiente de Correlação de Concordância Global (CCCG) [e.g. Barnhart02]. Este coeficiente tem sido aplicado em diversos âmbitos, designadamente no estudo da confiabilidade inter-observador [e.g. Heerink09], na determinação da qualidade de ajustamento de modelos de regressão [e.g. Bragada09] e na validação de métodos/instrumentos para medição de variáveis fisiológicas [e.g. Atkinson98].

## 2. METODOLOGIA

### 2.1 Materiais e equipamentos

#### 2.1.1 O sistema *ErgoVR*

O sistema *ErgoVR*, desenvolvido no Laboratório de Ergonomia da FMH/UTL [Teixeira, *in print*], visa permitir a recolha e análise de dados relativos ao comportamento humano durante a interacção com EVs.

O *ErgoVR* recorre a várias bibliotecas de acesso e uso gratuito, tendo como componentes principais: a) motor gráfico – responsável por apresentar o EV aos participantes, que pode ser criado recorrendo ao 3ds Max v2009, da Autodesk®, sendo, posteriormente, exportado através do programa OgreMax; b) motor de física – permite a detecção de colisões do “actor” (representação virtual do participante) com os objectos presentes no EV que podem revelar comportamentos de acordo com as leis da física; c) motor de som – permite que o participante seja exposto a uma multiplicidade de sons com características físicas realistas como, por exemplo, posicionamento, oclusão e obstrução da fonte sonora; d)

controlo de dispositivos – faz a ligação entre o *software* e o *hardware* (ex. sensores de movimento e interfaces de navegação) de forma a produzir uma resposta em tempo real entre os movimentos realizados pelo participante sobre o equipamento, e os movimentos realizados pelo “actor” dentro do EV; e) sistema de eventos – permite a colocação de gatilhos invisíveis no EV, que desencadeiam determinados eventos em resposta às acções do participante; f) sistema de registos – é o componente mais importante do sistema e grava os dados automaticamente. Os dados sobre a posição do “actor”, a orientação do olhar, eventos ocorridos durante a simulação (colisão com algum objecto, gatilho activado, detecção da escolha de direcção, se olhou para algum objecto específico, detecção de movimentos bruscos) são gravados com uma frequência aproximada de 60 vezes por segundo (valor configurável); g) visualizadores de registos – existem dois visualizadores de registos. Um que permite rever a simulação efectuada pelo participante através de um controlo do tipo gravador de vídeo, de três pontos de vista diferentes (o mesmo que o participante teve, um ponto de vista mais afastado, atrás do participante, e um ponto de vista de topo onde se observa o percurso realizado até ao momento). O outro visualizador apresenta os dados da simulação e extrapola algumas das variáveis pretendidas. Também é desenhado o percurso realizado pelo participante segundo um ponto de vista de topo.

Entre os principais *outputs* do *ErgoVR* encontram-se: tempos despendidos (total e parciais); distâncias percorridas (total e parciais); pausas (de duração variável); percursos e escolhas de rotas; eventos (ex. activação de gatilhos); informação sobre a posição do participante e direcção do olhar a cada instante. Os dados podem ser exportados por participante, ou por grupos de participantes, consoante as necessidades do estudo. No que se refere aos percursos, é possível gerar matrizes de exploração do espaço (gráficas ou numéricas), por participante ou grupo de participantes, indicando caminhos percorridos, áreas pisadas, entre outras opções.

### 2.1.2 Configuração do ErgoVR

A configuração do *ErgoVR*, usada neste estudo, no que se refere ao *hardware*, foi a seguinte: 2 sensores de movimento magnéticos da Ascension-Tech®, modelo Flock of Birds, com 6 DOF, para controlar os movimentos da cabeça (aplicado nos auscultadores) e da mão esquerda do participante (aplicado numa braçadeira de pulso); 1 *joystick* da Thrustmaster®, para navegação no mundo virtual; 1 *Head-Mounted Display* (HMD), da Sony®, modelo PLM-S700E, exibindo imagens com resolução de 800x600 a 32bits, com um ângulo de visão de cerca de 30° na diagonal e 28° na horizontal; auscultadores sem fios da Sony®, modelo MDR-RF800RK; 1 *eye-tracker* da Arrington Research®, modelo ViewPoint Monocular Laptop HMD Mountable; 1 *workstation* gráfica (com um processador Quad-core da Intel®, com 8GB de memória RAM e uma placa gráfica QuadroFX 4600 da nVIDIA®); 1 ecrã, que exibia uma cópia do EV, tal como visualizado pelo participante, para observação pelo investigador.

## 2.2 Envolvimento Virtual e cenário

O EV usado neste estudo consistia num espaço interior, num piso térreo, com características típicas de um edifício de escritórios, composto por duas grandes zonas: Zona 1 – salas e Zona 2 – escapatórias (ver Figura 1). A Zona 1 continha quatro salas (Sala de reuniões, Laboratório, Cafeteria e Armazém), cada uma com 144m<sup>2</sup> de área, nas quais os participantes deveriam entrar a fim de cumprir tarefas estipuladas. A Zona 2 consistia numa sequência de seis corredores em “T”, com 2 m de largura, sendo que apenas uma opção (esquerda/direita) tinha saída. O ambiente estava povoado por equipamento variado (ex. mesas, cadeiras, armários), de acordo com as tarefas que ocorreriam em cada sala, e diverso material de informação estava aplicado nas paredes (ex. painéis de mensagens, cartazes, sinalética e sinalização de segurança). Algumas das propriedades físicas do EV, tais como a luz/sombras, as texturas e os sons, foram manipuladas, com vista à obtenção de um grau de realismo aceitável que, em simultâneo, não compromettesse o funcionamento do sistema de RV (ex. provocar atrasos no refrescamento das imagens).

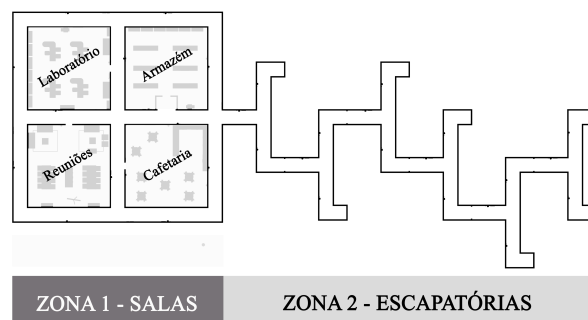
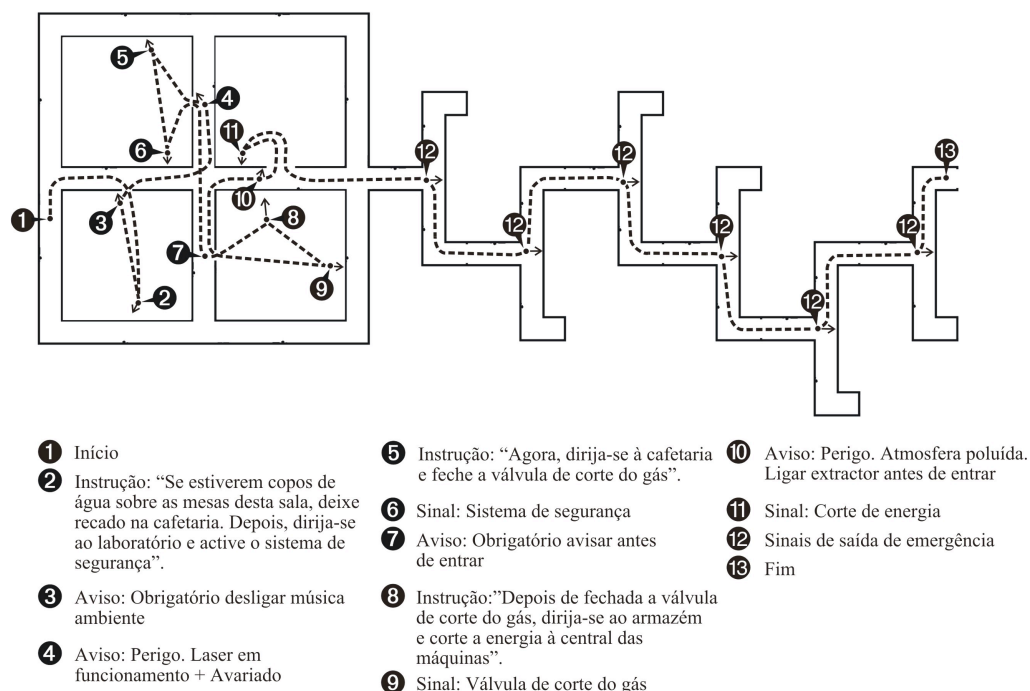


Figura 1. Planta do EV

O cenário criado para este estudo consistiu numa tarefa de inspecção de rotina, de final de dia, pré-encerramento das instalações que, a dado ponto, é interrompida por uma explosão, seguida de incêndio. Foi dito a cada participante que havia sido seleccionado para substituir um colega de trabalho, que se teria ausentado mais cedo devido a um problema de saúde. O participante estava, por isso, prestes a entrar numa secção da empresa onde nunca tinha ido antes, para desempenhar essa tarefa pela primeira vez. O colega ausente teria deixado instruções escritas em cada sala, indicando o que devia ser feito, com o intuito de facilitar a execução da tarefa. Assim, o participante deveria procurar as instruções, cumprir com o indicado e, só depois, passar para a sala seguinte. A primeira sala desta sequência era a Sala de reuniões. As instruções, escritas em quadros de parede, indicavam que deviam carregar em botões específicos (ex. cortar energia à central das máquinas), identificados por avisos/sinais de segurança e posicionados em diversos locais do EV. O esquema da Figura 2 ilustra a sequência pré-definida que o participante deveria seguir, o conteúdo das instruções e a localização dos elementos chave (botões/avisos e quadros de instruções) no EV. Na conversa com os participantes não foi feita qualquer referência à ocorrência da explosão/incêndio.



**Figura 2. Esquema ilustrando a sequência prevista para a tarefa a realizar no EV, com indicação das instruções e avisos, bem como da sua localização**

Os participantes foram informados que a simulação terminaria automaticamente se alcançassem a saída para o exterior (que se encontrava no final dos corredores que compõem a zona das escapatórias), ou se fosse atingido o tempo máximo (20 minutos) estabelecido para a realização da experiência, ou assim que manifestassem a intenção de abandonar a experiência.

### 2.3 Condições experimentais e estímulos

No estudo foram usadas amostras independentes, indexadas a cada uma das 5 condições experimentais, 4 delas resultantes da conjugação de 2 tipos de avisos de segurança (estáticos e dinâmicos) e de 2 tipos de ambiente (mais poluído e menos poluído) e uma condição de controlo ou neutra (sem avisos e com ambiente menos poluído). Em síntese, as condições experimentais foram designadas por: Estática Não Poluída (ENP), Dinâmica Não Poluída (DNP), Estática Poluída (EP), Dinâmica Poluída (DP) e Neutra (NE).

Os avisos de segurança eram constituídos por placas (30x40cm) com dois componentes, texto e imagem. No componente de texto estava inscrita uma mensagem, em português, que poderia ser de segurança (relativa à presença de um perigo nas imediações), de instrução (o que fazer) ou de informação (identificando um equipamento). No componente de imagem estavam aplicados símbolos de segurança, aplicados sobre um fundo colorido, de acordo com as recomendações da ISO 3864-1 [ISO02].

Os avisos estáticos e dinâmicos diferiam entre si quanto à modalidade e ao estado. Os estáticos eram apenas visuais, enquanto os dinâmicos eram visuais e auditivos (beep), auto-iluminados (2 vezes mais brilhante do que o fundo), complementados com 5 luzes intermitentes (4 flashes por segundo, com igual tempo de intervalo entre *on* e *off*), cor

de laranja, com 4 cm de diâmetro, posicionadas no topo e na base das placas. Os avisos dinâmicos tinham, ainda, dois estados possíveis, ligado e desligado, sendo activados por sensores de proximidade (gatilhos invisíveis). Exemplos de avisos estáticos e dinâmicos podem ser vistos na Figura 3.



**Figura 3. Imagens do EV, menos poluído, ilustrando um aviso "estático" (esq.) e o equivalente "dinâmico" (dir.)**

No EV estavam presentes 7 botões, com as respectivas placas de sinalização. Porém, estas não são todas equivalentes, podendo distinguir-se 2 grupos: a) os sinais: placas assinalando a presença de um equipamento, explicitamente referidas nas instruções e, por isso, expectavelmente procuradas de forma activa pelos participantes; b) avisos: placas alertando para um perigo potencial e requerendo uma acção específica para garantir a segurança, aplicadas na entrada das salas (no corredor junto à porta de entrada) ou no seu interior, nunca mencionadas nas instruções, muito embora fossem coerentes com o tipo de actividades praticada em cada sala. Relativamente às variáveis do design (ex. tipo de letra, organização, dimensões) os avisos e os sinais são idênticos.

Os dois tipos de ambiente, mais poluído e menos poluído, caracterizavam-se pela quantidade de informação visual



presente no EV (mais ou menos informação), pelo grau de legibilidade oferecido pelos avisos (degradados ou intactos) e pelo nível de iluminação do ambiente (adequado para tarefas com baixa exigência visual ou adequado para tarefas com elevada exigência visual). O tipo de ambiente foi pensado para facilitar, ou dificultar, a eficiência do desempenho dos participantes, avaliada pelo respeito pelos avisos/sinais (consonância comportamental), tempo associado à realização do pedido e distância percorrida no EV.

## 2.4 Amostra

Participaram neste estudo 150 indivíduos de ambos os sexos, com idades compreendidas entre os 18 e os 35 anos (média = 21,20 anos; DP = 2,85 anos). A cada uma das cinco condições experimentais, descritas na secção 2.3, foram aleatoriamente atribuídos 15 indivíduos de cada género. As amostras são homogéneas quanto à idade dos participantes. Todos os participantes eram alunos universitários, de áreas de estudo como o Desporto, a Reabilitação Psicomotora e o Design, entre outras, e declararam não ter tido qualquer experiência prévia com RV imersiva. Para participação no estudo foram considerados os seguintes critérios de elegibilidade: domínio da língua portuguesa (escrito e falado); no caso de ser mulher, não estar grávida; não sofrer de surdez nem ter algum tipo de deficiência na visão das cores e conseguir ler a informação exibida no EV. Adicionalmente, os participantes declararam estar em boas condições de saúde, quer físicas quer mentais, e não sofrer de alguma patologia que pudesse ser agravada pela participação no estudo.

## 2.5 Protocolo

O procedimento dividiu-se em três grandes momentos: 1) Actividades Preliminares: após receberem breves explicações sobre o estudo e os equipamentos, os participantes responderam a um questionário demográfico, preencheram o formulário de Consentimento Livre e Informado, realizaram o Teste de Ishihara [Ishihara88] para despiste de cegueira a cores e completaram uma sessão de treino, num EV específico para o efeito, na qual estava incluída a calibração dos equipamentos; 2) Sessão Experimental: os participantes foram indexados a uma das condições experimentais e, antes de iniciarem a simulação em RV, ouviram as instruções relativas ao cenário. Quando declararam estar prontos, deu-se início à simulação. Por questões de segurança, o investigador permaneceu, durante todo o tempo da simulação, na mesma sala com os participantes, que foram devidamente alertados de que podiam abandonar o procedimento a qualquer instante; 3) Questionário Pós-experimentação: após a sessão experimental os participantes responderam a um questionário para avaliar a sua experiência no EV.

## 3. RESULTADOS E DISCUSSÃO

O *ErgoVR* guardou uma sequência de registos num ficheiro, durante a simulação. Os dados foram processados após o final da simulação, para não ocupar os recursos de processamento. Para este estudo, cada registo

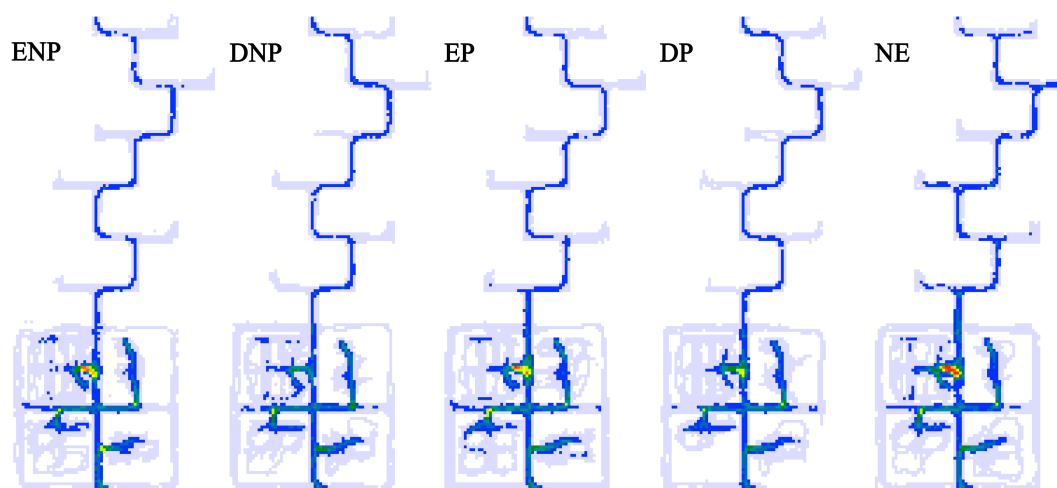
foi composto pela posição do “actor” no EV e pelo instante de tempo em que ocorreu o registo. Posteriormente, o visualizador de registos interpretou o ficheiro gerado e extrapolou diversos dados por participante e/ou grupo de participantes. Com esses dados foi possível gerar os percursos efectuados, bem como obter matrizes de exploração do espaço no EV, para as zonas de interesse.

A matriz de exploração espacial consiste numa grelha (com dimensões de quadrícula configuráveis), com uma cor e um número associado a cada quadrícula. O número representa a frequência com que cada quadrícula foi pisada durante a simulação. Neste estudo considerou-se como “quadrículas pisadas”, aquelas que foram intersectadas pelo ponto que representa o eixo central do “actor”. Cada vez que o “actor” entra na área da quadrícula a contagem da frequência de “pisadelas” aumenta uma unidade. No estudo efectuado, o valor apresentado na quadrícula é a soma das frequências de “pisadelas” de todos os indivíduos indexados à condição experimental que está a ser analisada. A cor de cada quadrícula é atribuída a partir de uma escala intervalar de 11 pseudo-cores predefinidas [Ware04], para indicar as zonas “quentes” (cores quentes = mais pisadas) e “frias” (cores frias = menos pisadas). O valor mais alto de todas as quadrículas (quadrícula mais “pisada”), considerando as 5 condições experimentais, serve como base para a normalização das cores atribuídas. Em alternativa, a apresentação dos dados pode ser disponibilizada numa versão numérica, onde a matriz com as frequências de “pisadelas” é disponibilizada em vectores.

Na Figura 4 é possível visualizar as 5 matrizes (da totalidade do EV), relativas a cada uma das condições experimentais, com quadrículas coloridas de dimensão 50 por 50 cm. Esta dimensão foi adoptada em virtude de se aproximar do valor do percentil 95 da largura dos ombros do sexo masculino (51 cm) [Pheasant03].

A matriz tem o seu ponto de origem no canto superior esquerdo da janela onde é possível visualizar a planta do EV e, por omissão, inclui todo o espaço virtual. No entanto, para facilitar a análise, também é possível criar áreas de exclusão (servem para excluir zonas como, por exemplo, locais onde os participantes nunca poderiam navegar), áreas de inclusão (servem para permitir exportar dados apenas de uma zona específica) e áreas para criação de gatilhos manuais (permitem criar zonas no EV que servem para detecção da passagem do participante, permitindo assim agrupar os dados por zonas).

Uma vez que havia interesse na avaliação da exploração espacial efectuada por zonas específicas, foram seleccionadas 5 áreas de inclusão (às quais foram excluídas todas as quadrículas não pisáveis). Essas áreas correspondem a cada uma das quatro salas, incluindo o corredor adjacente à parede que tem a porta de entrada, e a zona das escapatórias.



**Figura 4. Matrizes de exploração do EV, por condição experimental**

Para cada uma das zonas, a avaliação do “acordo” da exploração espacial nas 5 condições experimentais, foi efectuada através da aplicação do CCCG [Barnhart02] sobre as frequências com que as quadrículas foram pisadas nas 5 condições experimentais. Os valores do CCCG obtidos são apresentados na primeira linha da Tabela 1. Atendendo a que este coeficiente toma valores no intervalo  $[-1,1]$ , podemos verificar que a concordância é bastante elevada, em todas as zonas analisadas, o que permite afirmar que a exploração do espaço não foi fortemente influenciada pela condição experimental. Ainda assim refira-se que os valores mais baixos do CCC foram encontrados nas Escapatórias e no Armazém, e os mais elevados na Cafeteria e na Sala de reuniões.

	Arm.	Caf.	Esc.	Lab.	Reu.
50 por 50 cm	0,843	0,948	0,838	0,918	0,946
75 por 75 cm	0,856	0,960	0,841	0,927	0,961
100 por 100 cm	0,858	0,972	0,860	0,940	0,966

**Tabela 1 – Valores do CCCG por zona**

Os valores do CCC parcial, isto é, os valores do CCC considerando as condições experimentais duas a duas, foram também calculados. No Armazém, os valores mais elevados de CCC foram obtidos entre as condições experimentais ENP/N, EP/N, ENP/EP, e os valores mais baixos de CCC foram registados entre as condições DNP/N, DNP/EP (ver Tabela 2). Nas Escapatórias os valores mais elevados registaram-se nos casos DNP/DP, ENP/DNP, e os mais baixos nos casos DP/N, DNP/N (ver Tabela 3). Os valores do CCC parcial obtidos para as zonas do Laboratório (ver Tabela 4), da Cafeteria (ver Tabela 5) e da Sala de reuniões (ver Tabela 6) foram muito idênticos.

Uma vez que o visualizador de registos do *ErgoVR* permite a geração de matrizes com diferentes dimensões de quadrícula, o CCCG foi também aplicado aos valores obtidos em matrizes com quadrículas de 75 e 100 cm de largura. A alteração da dimensão da quadrícula não originou alterações relevantes nos valores do CCCG. Destaca-se o facto dos valores do CCCG diminuírem com a diminuição da largura da quadrícula. Um resumo dos

valores de CCCG, para todas as dimensões de quadrícula consideradas, pode ser visualizado na Tabela 1.

	DNP	EP	DP	N
ENP	0,762	0,909	0,877	0,926
DNP		0,717	0,838	0,658
EP			0,884	0,925
DP				0,833

**Tabela 2 – Valores do CCC parcial para o Armazém**

	DNP	EP	DP	N
ENP	0,895	0,889	0,887	0,825
DNP		0,873	0,909	0,751
EP			0,852	0,823
DP				0,720

**Tabela 3 – Valores do CCC parcial para as Escapatórias**

	DNP	EP	DP	N
ENP	0,945	0,921	0,917	0,952
DNP		0,892	0,927	0,945
EP			0,871	0,903
DP				0,909

**Tabela 4 – Valores do CCC parcial para o Laboratório**

	DNP	EP	DP	N
ENP	0,941	0,960	0,951	0,959
DNP		0,937	0,942	0,944
EP			0,955	0,947
DP				0,942

**Tabela 5 – Valores do CCC parcial para a Cafeteria**

	DNP	EP	DP	N
ENP	0,929	0,935	0,942	0,950
DNP		0,947	0,955	0,943
EP			0,952	0,950
DP				0,959

**Tabela 6 – Valores do CCC parcial para a Sala de reuniões**

#### 4. CONCLUSÃO

O modo como os participantes numa experiência de RV exploram o espaço pode ser usado para avaliar a sua interacção com EVs. Neste contexto, este estudo avaliou a interacção com EVs, em função de diferentes condições experimentais, recorrendo a matrizes de exploração do espaço, geradas com base nos dados recolhidos, de forma automática, pelo sistema *ErgoVR*.

As matrizes contêm indicação do número de vezes que cada quadrícula é pisada e espelham o modo como os participantes exploraram o EV nas cinco condições experimentais. Partindo do pressuposto que os avisos mais salientes (dinâmicos) seriam detectados mais facilmente, foi colocada a hipótese de que essas diferenças estariam reflectidas na exploração espacial do EV. Esperava-se que nas condições experimentais onde os avisos são menos salientes (estáticos) e o ambiente se apresenta mais poluído, houvesse uma maior deambulação do participante no EV, originando uma maior dispersão quanto às quadrículas pisadas, bem como frequências mais elevadas de “pisadelas” em algumas quadrículas.

O Coeficiente de Correlação de Concordância foi usado para determinar o grau de “acordo”, entre as condições experimentais, quanto à frequência com que as quadrículas foram pisadas. Os resultados obtidos revelam valores bastante elevados de “acordo” entre as matrizes, o que permite concluir que não houve grande influência das condições experimentais sobre a exploração do espaço.

Neste estudo não foi considerado o comportamento individual dos participantes na experiência de RV, uma vez que o coeficiente foi utilizado para avaliar a concordância da frequência com que as quadrículas da matriz são pisadas pelos 30 participantes que foram atribuídos a cada uma das 5 condições experimentais. Uma limitação deste estudo é o facto de, para cada uma das condições experimentais, não se ter tido em conta se a frequência registada para uma determinada quadrícula se deve a um único indivíduo, que pisou repetidamente a mesma quadrícula, ou se foram diversos indivíduos a pisar a mesma quadrícula. Em estudos futuros pretende-se avaliar os percursos individuais, comparar os percursos efectuados com percursos de referência (mais curtos, mas rápidos, mais fáceis) e, através da atribuição de pontuações ao grau de afastamento face à referência, obter valores de desempenho. Informações adicionais, como a orientação do corpo do “actor” e o seu campo de visão, a cada instante da interacção, poderão complementar as análises.

O uso da RV, neste tipo de estudos, permite a fácil obtenção de informação referente aos percursos efectuados pelos participantes, assim como os tempos e as distâncias. Este tipo de informação seria mais difícil de obter em estudos de campo. A RV garante, em simultâneo, uma óptima relação entre o controlo rigoroso das variáveis e o necessário grau de validade ecológica do estudo.

#### 5. AGRADECIMENTOS

Este trabalho foi financiado por uma Bolsa de Doutorado atribuída pela FCT – Fundação para a Ciência e a Tecnologia (SFRH/BD/21662/2005) e foi desenvolvido

no âmbito de um Projecto de I&D financiado pela FCT (PTDC-PSI-69462-2006).

#### 6. REFERÊNCIAS

- [Anderson00] Anderson, J.R. (2000). *Cognitive psychology and its implications*. 5th Edition. New York: Worth Publishers.
- [Atkinson98] Atkinson, G., Nevill, A. (1997). Comment on the use of concordance correlation to assess the agreement between two variables. *Biometrics*, 53, 775-777.
- [Barnhart02] Barnhart, H.X., Haber, M., Song, J.L. (2002). Overall concordance correlation coefficient for evaluating agreement among multiple observers. *Biometrics*, 58, 1020-1027.
- [Bragada09] Bragada, J.A., Magalhães, P.P., Vasques, C.S., Barbosa, T.M., Lopes, V.P. (2009). Net heart rate to prescribe physical activity in middle-aged to older active adults. *Journal of Sports Science and Medicine*, 8, 616-621.
- [Carrasco05] Carrasco, J.L., Jover, L. (2005). Concordance correlation coefficient applied to discrete data. *Statistics in Medicine*, 24, 4021-4034.
- [Duarte, in print] Duarte, M.E.C., Rebelo, F., Wogalter, M.S. (in print). Virtual reality (VR) and its potential for evaluating warning compliance. *Human Factors in Manufacturing & Service Industries*.
- [Gamberini03] Gamberini, L., Cottone, P., Spagnolli, A., Varotto, D., Mantovani, G. (2003). Responding to a fire emergency in a virtual environment: different patterns of action for different situations. *Ergonomics*, 46(8), 842-858.
- [Glover97] Glover, B.L., Wogalter, M.S. (1997). Using a computer simulated world to study behavioral compliance with warnings: Effects of salience and gender. *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (Vol. 1 and 2, pp. 1283-1287). Santa Monica: Human Factors and Ergonomics Soc.
- [Heerink09] Heerink, M., Kröse, B., Evers, V., Wielinga, B. (2009). Relating conversational expressiveness to social presence and acceptance of an assistive social robot. *Virtual Reality*, 14, 77-84.
- [Ishihara88] Ishihara, S. (1988). *Test for Colour-Blindness* (38th ed.). Tokyo: Kanehara & Co., Ltd.
- [ISO02] International Organization for Standardization (2002). *Graphical Symbols – Safety Colors and Safety Signs. Part 1: Design Principles for safety Signs in Workplaces and Public Areas, ISO 3864-1*. Geneve, Switzerland: International Organization for Standardization.
- [Kalsher06] Kalsher, M.J., Williams, K.J. (2006). Behavioral Compliance: Theory, Methodology, and Results. In M.S. Wogalter (Ed.), *Handbook of warnings* (pp. 313-332). Mahwah: Lawrence Erlbaum Associates, Inc.

- [Lin89] Lin, L.I.-K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 45, 255-268.
- [Lin02] Lin, L., Hedayat, A.S., Sinha, B., Yang, M. (2002). Statistical methods in assessing agreement: models, issues, and tools. *Journal of American Statistical Association*, 97(457), 257-270.
- [Pheasant03] Pheasant, S. (2003). *Bodyspace: Anthropometry, Ergonomics and the Design of Work*. 2nd Edition. Boca Raton: Taylor & Francis.
- [Ren08] Ren, A., Chen, C., Luo, Y. (2008). Simulation of Emergency Evacuation in Virtual Reality. *Tsinghua Science & Technology*, 13(5), 674-680.
- [Silver99] Silver, N.C., Braun, C.C. (1999). Behavior. In M.S. Wogalter, D.M. Dejoy, K.R. Laughery (Eds.), *Warnings and Risk Communication* (pp. 245-262). London: Taylor & Francis Ltd.
- [Tang09] Tang, C.-H., Wu, W.-T., Lin, C.-Y. (2009). Using virtual reality to determine how emergency signs facilitate way-finding. *Applied Ergonomics*, 40(4), 722-730.
- [Teixeira, in print] Teixeira, L., Rebelo, F., Filgueiras, E. (in print). Human Interaction data acquisition software for Virtual Reality: A User-Centered Design approach. *AHFE 2010, 3rd International Conference on Applied Human Factors and Ergonomics*. 17-20 July 2010, Miami, Florida, USA.
- [Ware04] Ware, C. (2004). *Information Visualization: Perception for Design*. 2nd Edition. San Francisco: Morgan Kaufmann Publishers.
- [Wogalter87] Wogalter, M.S., Godfrey, S.S., Fontenelle, G.A., Desaulniers, D.R., Rothstein, P.R., Laughery, K.R. (1987). Effectiveness of Warnings. *Human Factors*, 29(5), 599-612.
- [Wogalter93] Wogalter, M.S., Kalsher, M.J., Racicot, B.M. (1993). Behavioral compliance with warnings: Effects of voice, context, and location. *Safety Science*, 16(5-6), 637-654.
- [Wogalter06] Wogalter, M.S., Vigilante, J.W.J. (2006). Attention Switch and Maintenance. In M.S. Wogalter (Ed.), *Handbook of warnings* (pp. 245-266). Mahwah: Lawrence Erlbaum Associates, Inc.

# Prototipagem rápida de ambientes ubíquos

José Luís Silva      Óscar R. Ribeiro      João M. Fernandes  
 José Creissac Campos  
 Dep. Informática / CCTC, Universidade do Minho, Braga, Portugal  
 {jlsilva, orribeiro, jmf, jose.campos}@di.uminho.pt

Michael D. Harrison  
 Newcastle University, United Kingdom  
 michael.harrison@ncl.ac.uk

## Resumo

*A experiência de utilização de ambiente ubíquos é um factor determinante no seu sucesso. As características de tais sistemas devem ser exploradas o mais cedo possível para antecipar potenciais problemas de utilização por parte do utilizador e para reduzir custos de re-design. No entanto, o desenvolvimento antecipado de protótipos a serem avaliados no ambiente final pode ser disruptivo e tornar-se inaceitável. O desenvolvimento de protótipos de ambientes ubíquos pode ajudar, fornecendo indicações de como o utilizador irá reagir perante os ambientes. Este artigo descreve o APEX, uma plataforma de prototipagem rápida de ambientes ubíquos que junta a CPN Tools com um servidor de aplicações 3D existente. Os protótipos desenvolvidos com o APEX permitem que os utilizadores naveguem num mundo virtual, podendo experimentar muitas das características do design proposto. A arquitectura do APEX e a modelação baseada em CPN são descritas. Um exemplo ilustra a abordagem.*

## Palavras-Chave

*Computação ubíqua, modelação, ambientes virtuais*

## 1. Introdução

A computação ubíqua impõe novos desafios de usabilidade para os *designers* e *developers* de sistemas interactivos. Dado que nestes sistemas os utilizadores estão imersos, o efeito que os sistemas tem sobre a experiência de utilização de um dado espaço torna-se um elemento importante que contribui para o sucesso do *design*. Um sistema bem sucedido deve, por exemplo, conseguir transformar um ambiente estéril num lugar no qual as pessoas estejam em harmonia com o ambiente e seus propósitos. Um exemplo de criação de tal experiência num ambiente ubíquo pode ser o de efectuar um *check-in* num aeroporto e mover-se com o mínimo de atraso através das várias etapas da partida. Outro exemplo pode ser um ambiente que fornece informação adaptada aos utilizadores de uma biblioteca, facilitando-lhes localizar a prateleira onde o livro desejado se encontra. A noção de *experiência* é difícil de capturar como um requisito que pode ser demonstrado num sistema. Actualmente, não existem técnicas que possam ser utilizadas permitindo analisar especificações contra diferentes noções de experiência (ver [Harrison 08] para uma discussão deste tema).

Este artigo aborda a prototipagem de ambientes ubíquos projectados para melhorar ambientes físicos reais através da utilização de sensores, ecrãs públicos e dispositivos pes-

soais. Um aspecto particularmente relevante destes sistemas é a forma como os utilizadores interagem com o ambiente, quer através de interações explícitas com o sistema, quer de interações implícitas que surgem através de mudanças do contexto. Aqui o contexto pode incluir a posição ou as diferentes etapas que tem de ser cumpridas pelo utilizador por forma a atingir um determinado objectivo (por exemplo: *check-in*, controlo do passaporte, digitalização do bilhete de embarque).

O objectivo do trabalho é possibilitar a avaliação de ambientes ubíquos precocemente, antes mesmo da sua efectiva instalação *in situ*. O artigo é baseado no pressuposto que protótipos serão utilizados para explorar o impacto que o *design* de um dado ambiente ubíquo terá nos utilizadores consoante eles se mexem, interagem e acedem a serviços dentro desse ambiente. Para evitar custos de desenvolvimento desnecessários, o *design* é a explorado desde cedo através de protótipos baseados em modelos, recorrendo a ambientes virtuais. O artigo descreve uma plataforma de prototipagem (APEX) que utiliza modelos de redes de Petri coloridas [Jensen 07] (CPN - Coloured Petri Nets). O APEX interliga um modelo CPN com um servidor de aplicações 3D (OpenSimulator<sup>1</sup>). A plataforma permite a prototipagem de ambientes ubíquos, possibilitando que os

<sup>1</sup><http://opensimulator.org> (acedido em: 14 de Junho de 2010)

utilizadores naveguem na simulação do ambiente virtual através do controlo (actualmente através do rato e do teclado) de um avatar que o representa, tornando possível avaliar problemas de usabilidade, incluindo noções de experiência de utilização.

A estrutura do artigo é a seguinte. A secção 2 apresenta trabalhos relacionados presentes na literatura. A secção 3 descreve a arquitectura do APEX. A plataforma é ilustrada através de uma biblioteca inteligente que detecta a presença dos seus utilizadores e os guia até as prateleiras onde se encontram os livros que estes pretendem requisitar. A secção 4 descreve como o exemplo é modelado. A secção 5 descreve a utilização da plataforma e a secção 6 apresenta conclusões e trabalho futuro.

## 2. Literatura relacionada

Apesar de avanços consideráveis no desenvolvimento de sistemas ubíquos, continua a existir uma tendência para que o seu desenvolvimento e avaliação sejam feitos principalmente através de experimentação, desenvolvendo protótipos de dispositivos a avaliar em sistemas parcialmente desenvolvidos (ver [Davies 05] para uma discussão deste tema). A questão de como avaliar os sistemas num contexto real permanece um tópico em aberto [Abowd 05]. Outro importante aspecto de avaliação é saber como explorar a experiência criada por um dado sistema. Nesse campo, existe uma literatura substancial proveniente do campo do *design* (ver, por exemplo, [Buchenau 00]). Um exemplo de uma abordagem típica no campo do *design* é desenvolver protótipos não-funcionais que potenciais utilizadores podem transportar no contexto onde o sistema real irá ser utilizado. O objectivo é obter informações acerca do modo como o *design* proposto irá ser experimentado.

O APEX é projectado para satisfazer 3 requisitos. O primeiro é permitir o desenvolvimento rápido de protótipos. Embora existam várias plataformas de computação ubíqua (e.g., [Braubach 02, Garlan 02, Harter 01]), o desenvolvimento de protótipos que permitam avaliar a reacção dos utilizadores a ambientes ubíquos é uma área ainda relativamente pouco explorada.

O segundo requisito é a utilização de ambientes 3D para construir simulações que podem ser exploradas de forma realística por utilizadores. Servidores de aplicações 3D, tais como o SecondLife<sup>2</sup> ou o OpenSimulator fornecem uma via de desenvolvimento rápido de mundos virtuais. O OpenSimulator em particular tem a vantagem de disponibilizar o código fonte, o que significa que pode ser programado permitindo uma maior configurabilidade e extensibilidade.

O terceiro requisito é possuir uma abordagem para a modelação de sistemas ubíquos. As redes de Petri constituem uma linguagem de modelação gráfica expressiva e tem sido utilizada para descrever ambientes virtuais. Abordagens anteriores baseadas em redes de Petri incluem: Hybrid high level Nets (HyNets) [Massink 99], Flow-nets [Smith 99], Interactive Cooperative Objects (ICO)

[Navarre 05] e CPNs [Jensen 07].

Um objectivo desta plataforma é integrar a abordagem de modelação com abordagens analíticas que permitam validar propriedades nos ambientes ubíquos que sejam relevantes aquando da sua utilização. Neste contexto, a utilização de CPNs surgiu como escolha lógica dado o bom suporte ao nível de ferramentas de animação e análise de modelos.

Embora possam ser encontradas na literatura várias abordagens à prototipagem de computação ubíqua, somente um número limitado refere aspectos de interacção do utilizador com o ambiente e da experiência obtida pelo utilizador. Alguns sistemas (e.g. [Li 04]) recorrem às técnicas de *Wizard of Oz* para evitar a instalação de sensores no ambiente. Outros, tais como 3DSim [Shirehjini 05], UbiWorld [Disz 97] ou o trabalho de O'Neill et al. [O'Neill 09] têm visões similares à nossa.

A abordagem de O'Neill et al. é a que mais se aproxima, utilizando modelos e um motor de jogos 3D para a prototipagem de ambientes ubíquos (3DSim e UbiWorld assemelham-se mais a *frameworks* de programação). Acreditamos que o uso de servidores de aplicações 3D (OpenSimulator) tem algumas vantagens comparativamente com a utilização de um motor de jogos. Os servidores de aplicações 3D suportam a criação de ambientes virtuais em tempo real utilizando ferramentas de construção de objectos e são facilmente extensíveis através do carregamento de módulos. No caso dos motores de jogos, o ambiente deve ser previamente criado utilizando um editor de mapas. Utilizar um servidor de aplicações 3D significa que a abordagem é flexível. Uma variedade de clientes, personalizáveis em aparência, podem aceder o mundo virtual em simultâneo utilizando diferentes protocolos e o desenvolvimento dos ambientes pode ser efectuado através de linguagens de programação variadas.

## 3. Arquitectura

O APEX utiliza a CPN Tools<sup>3</sup> para modelar o comportamento de ambientes virtuais. Esse comportamento é então ligado ao servidor de aplicações 3D (OpenSimulator). Modelos CPN dos diferentes tipos de dispositivos presentes no ambiente (e.g. sensores, ecrãs, dispositivos pessoais) devem ser criados. O objectivo é desenvolver um estilo genérico de CPN relevante para a modelação de ambiente virtuais, incluindo modelos que podem ser instanciados no espaço real no qual o sistema foi definido para operar.

O OpenSimulator permite a criação interactiva de ambientes virtuais suficientemente ricos permitindo aos utilizadores visualizar as características dos sistemas reais que estes representam. Ambientes e dispositivos pré-definidos podem ser utilizados nesse processo de criação.

Uma vez o modelo CPN e o ambiente criados, um componente do APEX é responsável por ligá-los. Vários utilizadores podem estabelecer ligações à simulação com diferentes pontos de vista no servidor OpenSimulator. Os

<sup>2</sup><http://secondlife.com> (acedido em: 14 de Junho de 2010)

<sup>3</sup><http://wiki.daimi.au.dk/cpntools/> (acedido em: 14 de Junho de 2010)

utilizadores podem navegar e interagir com a simulação do ambiente ubíquo, permitindo avaliar questões relacionadas com o design proposto.

A visão global da arquitectura da plataforma é apresentada na figura 1. Esta é composta por 3 partes:

1. um *componente comportamental* responsável por gerir o comportamento do protótipo, incluindo a descrição, análise e validação do comportamento do ambiente virtual;
2. um *componente arquitectural* responsável por gerir a aparência física do protótipo, incluindo a gestão da simulação 3D e da construção do ambiente virtual;
3. um *componente de comunicação/execução* responsável pela troca de dados entre todos os componentes da plataforma.

### 3.1. Componente comportamental

Este componente utiliza a linguagem de modelação CPN para descrever o comportamento do ambiente virtual em resposta a acções do utilizador e a mudanças do contexto. Um modelo CPN genérico base foi desenvolvido contendo:

1. um módulo para inicializar a simulação e para estabelecer a conexão entre o modelo CPN representado na CPN Tools e o OpenSimulator;
2. um módulo que recebe dados do utilizador (por exemplo, identidade e posição) a partir do OpenSimulator quando este se move, utilizando-os para actualizar *tokens* apropriados do modelo CPN;
3. um módulo descrevendo o comportamento de cada um dos tipos de dispositivos do sistema.

Um exemplo de um modelo CPN para o *componente comportamental* é apresentado na secção 4.

### 3.2. Componente arquitectural

Este componente utiliza o OpenSimulator para definir características 3D da simulação apresentada aos utilizadores e permite a navegação destes na simulação. As características 3D da simulação incluem a localização, o aspecto visual e a física de cada objecto no ambiente. Estas definições são efectuadas utilizando uma ferramenta de visualização, e.g. o Hippo Viewer<sup>4</sup> ou o Linden Lab's Second Life viewer<sup>5</sup>. Existem outros visualizadores compatíveis disponíveis (ver <http://opensimulator.org/wiki/Connecting>). Actualmente alguns desses visualizadores permitem somente a exploração do ambiente sem fornecer qualquer ferramenta de edição.

Este componente é também responsável por permitir a navegação livre e a interacção no ambiente. A interacção

pode ser efectuada explicitamente pelos utilizadores utilizando dispositivos (virtuais) e implicitamente através de mudanças do contexto. A plataforma permite a conexão de vários utilizadores de, possivelmente, diferentes localizações ao mesmo ambiente virtual através da Web. Ambientes e objectos pré-definidos podem ser guardados/carregados utilizando ficheiros OAR (Opensim Archive files). Todas as entidades (objectos, terrenos, texturas, etc.) são codificadas nesses ficheiros no formato utilizado pelo OpenSimulator para guardar dados em ficheiros.

### 3.3. Componente de comunicação/execução

Este componente é uma DLL (*dynamic-link library*) responsável pelo carregamento do ambiente ubíquo e por utilizar os modelos CPN para o dirigir. A comunicação é baseada no Comms/CPN [Gallasch 01], uma biblioteca desenvolvida para conectar a CPN Tools com processos Java e C. Como os módulos para o OpenSimulator são desenvolvidos em C#, um novo pacote de comunicação C#/CPN teve de ser desenvolvido. Este pacote envia informações para a CPN Tools quando mudanças no ambiente ocorrem e é também responsável por modificar o ambiente em resposta a dados enviados pela CPN Tools. Adicionalmente, este componente controla o carregamento/gravação de objectos e ambientes no OpenSimulator.

### 4. Exemplo: modelação com CPNs

O exemplo utilizado para ilustrar o sistema é uma biblioteca inteligente. Todos os livros são identificados com etiquetas RFID e são dispostos em prateleiras que tem LEDs associadas. São utilizados ecrãs para fornecer informação aos utilizadores da biblioteca. Um utilizador registado pode entrar/sair da biblioteca através das portas.

Quando um utilizador registado chega à porta de entrada, esta abre e o ecrã associado apresenta os livros que foram previamente requisitados por ele (e.g. através da página Web da biblioteca). De seguida o sistema guia, com informação em ecrãs presentes na biblioteca, o utilizador até aos livros requisitados através do uso de sensores que reconhecem a posição do utilizador em tempo real. Assim que os utilizadores se aproximam da localização dos referidos livros (distância configurável no modelo CPN), uma das luzes com uma cor específica liga-se. Uma vez que vários utilizadores podem estar à procura de livros em localizações próximas, este método facilita a distinção dos diferentes pedidos, dado que cada utilizador utilizará uma cor diferente, previamente indicada quando este chega à biblioteca. Quando o utilizador se dirigir para a porta de saída, uma lista personalizada dos livros requisitados será apresentada no ecrã perto da porta e esta será aberta caso o utilizador possa requisitar os livros. À saída do utilizador, os RFIDs dos livros são detectados e consequentemente requisitados automaticamente.

#### 4.1. Inicializando a simulação

As condições iniciais da simulação estão definidas no módulo CPN apresentado na figura 2. A execução da transição “*initialise simulation*” inicializa

<sup>4</sup><http://mjm-labs.com/viewer/> (acedido em: 14 Junho de 2010)

<sup>5</sup><http://secondlife.com/support/downloads> (acedido em: 14 Junho de 2010)



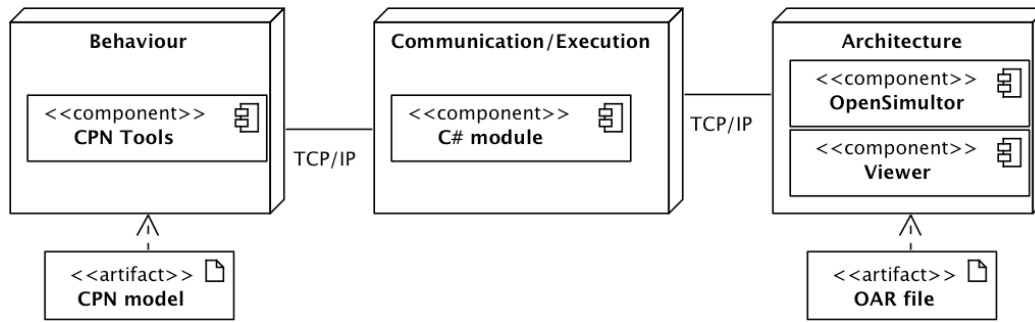


Figura 1. Arquitectura da plataforma APEX

a configuração inicial da simulação. Neste caso a configuração inclui 3 lugares: “users”, “gates” e “bookshelves”. As *Etiquetas de fusão* (e.g., etiqueta books no lugar bookshelves) permitem que instâncias desses lugares apareçam em outras partes do modelo CPN. Esses lugares são chamados *fusion places*. A transição “initialise simulation” cria uma nova conexão e utiliza dois lugares para controlar a execução do modelo CPN: “init” para limitar a execução da transição a uma ocorrência e “run” para informar os outros módulos CPN que a simulação está em execução.

#### 4.2. Leitura das posições do utilizador

A figura 3 apresenta o módulo CPN responsável por colectar os dados dos utilizadores. A transição “read user id” lê o identificador de um utilizador. Um *token* com o identificador do utilizador associado é introduzido no lugar “read user ids”. Este é utilizado para ler a nova posição do utilizador através da transição “read and update user position” que também actualiza o *token* do respectivo utilizador. A expressão “isThisUser(u, uId)” na guarda desta transição garante que o *token* do utilizador que é actualizado corresponde ao identificador que foi previamente lido. Neste modelo, o número de utilizadores permanece constante durante cada sessão de uma simulação, i.e. os utilizadores que se podem conectar à simulação em tempo de execução devem estar previamente inseridos no modelo (variável “initial USERS”).

Os módulos CPN leitores de posições dos utilizadores provenientes da simulação e os que descrevem o comportamento dos dispositivos executam concorrentemente. A precedência das transições dos dispositivos sobre a aquisição de dados é garantida através da guarda “not (hadASignificantMovement(u))” sobre a transição “read and update user position”. Um movimento de um utilizador é considerado significativo (para um dispositivo) quando a nova posição está “próxima” do dispositivo. Fica desta forma garantido que não ocorrem comportamentos indevidos devido a leituras antecipadas.

#### 4.3. Modelação dos dispositivos do sistema

Cada tipo de objecto dinâmico (dispositivo) presente no ambiente ubíquo simulado precisa de um correspondente

módulo CPN que descreva o seu comportamento. Os *fusion places* são a base do processo de criação desses módulos. Para criar os modelos desses módulos os *fusion places* necessários provenientes do modelo da figura 2 devem ser clonados nesses modelos. Neste exemplo os lugares “users”, “gates” e “bookshelves” contêm os *tokens* relativos aos objectos dinâmicos presentes. Estes *tokens* são utilizados para modelar o comportamento do sistema. Isto é efectuado através da combinação de *fusion places*, lugares normais, transições, funções (descritas na linguagem CPN ML associada à CPN Tools) e condições. As transições tem um papel importante nesse processo já que são elas que estabelecem a conexão entre os modelos CPN e o simulador, através da utilização de funções CPN ML. As funções também são utilizadas para descrever algum comportamento que não seja expressado estruturalmente pela rede de Petri. Um exemplo do modelo de um desses dispositivos é o modelo da porta de entrada apresentado na figura 4.

A porta de entrada é equipada com um sensor que reconhece quando um utilizador se aproxima para entrar na biblioteca. A transição “show info and open gate” representa a acção da porta de entrada. Esta mostra os livros requisitados no ecrã e abre a porta de entrada. Estas acções ocorrem quando o sensor da porta detecta a chegada de um utilizador (modelado com a condição “isArrivingToGateArea(u, g)”). Quando a porta está aberta e outro utilizador registado entra na zona da porta, a transição “add a user” é executada incluindo o utilizador no conjunto de utilizadores perto da referida porta.

Quando a transição “show default and close gate” é executada, informação por omissão é mostrada no ecrã e a porta é fechada. Para que isto aconteça um utilizador tem de se afastar da porta e mais nenhum utilizador pode estar perto dela. No caso de outros utilizadores estarem perto da porta, a transição “remove a user and update info” remove o referido utilizador do conjunto de utilizadores que estão perto da porta. Caso existisse informação para esse utilizador no ecrã, essa é trocada por informação relevante para outro utilizador que ainda esteja perto da porta.

Utilizando a ferramenta *State Space tool* fornecida com a CPN Tools podem validar-se propriedades dos modelos.

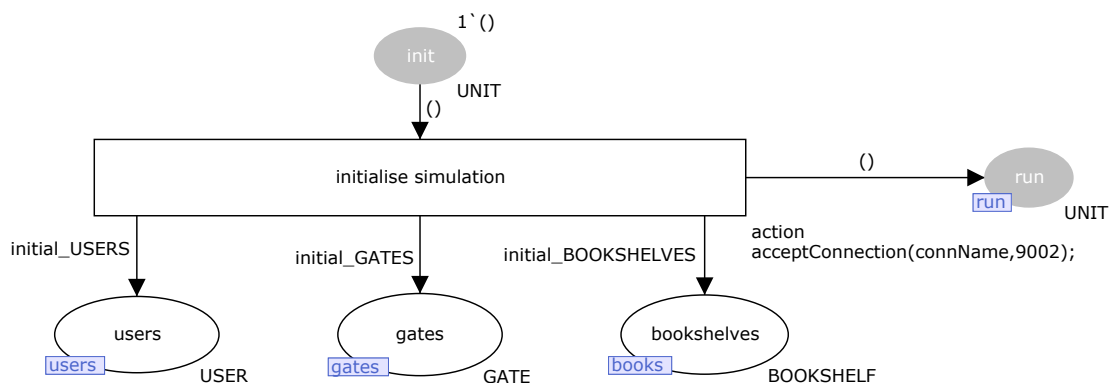


Figura 2. Módulo CPN para inicializar a simulação da biblioteca.

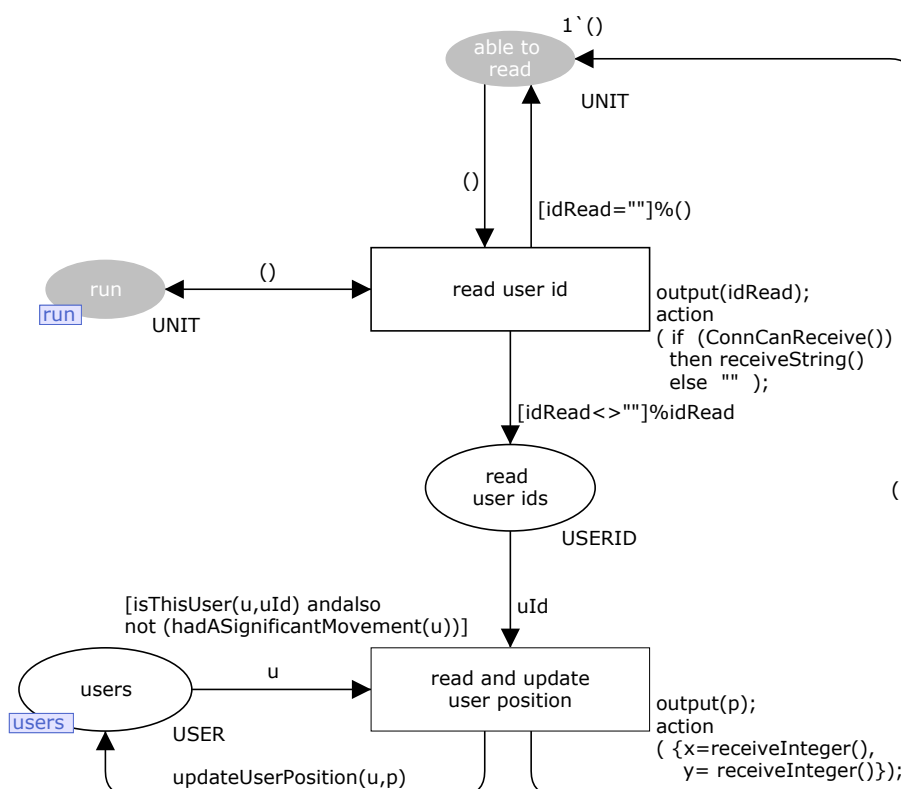


Figura 3. Módulo CPN para aquisição de dados dos utilizadores.

Por exemplo, a propriedade de acessibilidade, i.e. se todos os lugares são acessíveis ou se um dado lugar é acessível a partir de outro lugar numa dada situação, são exemplos de propriedades que podem ser facilmente validadas através da utilização de funções fornecidas pela *State Space tool* (c.f. *AllReachable()*, *Reachable(node,node)*).

## 5. Utilização

Como foi referido na secção 1, o objectivo do APEX é suportar o *design* e análise de sistemas ubíquos. A pessoa responsável pelo desenvolvimento do protótipo cria modelos CPN, tal como ilustrado na secção 4. Dependendo da necessidade de novos tipos de dispositivos, poderá ter que desenvolver novos modelos ou apenas reutilizar modelos existentes.

Cada dispositivo e cada utilizador é representado no modelo CPN por um *token* inserido no respectivo lugar etiquetado como *lugar de fusão* (e.g. veja o lugar *users* na figura 4). Cada um desses *tokens* tem um identificador que é comum aos identificadores dos objectos presentes na simulação. Estes dados são usados pelas funções CPN ML conjuntamente com instruções (e.g. *open*, *close*) para indicar mudanças que devem ser reflectidas no OpenSimulator. Estas indicações são responsáveis por reflectir mudanças nos objectos do ambiente de acordo com o estado do modelo CPN. Como exemplo, a figura 5 mostra um pedaço de código que procura objectos que devem ser modificados, de acordo com instruções recebidas do modelo CPN, e efectua as respectivas alterações. Nesse pedaço de código, a posição da porta é alterada de acordo com as acções *open*

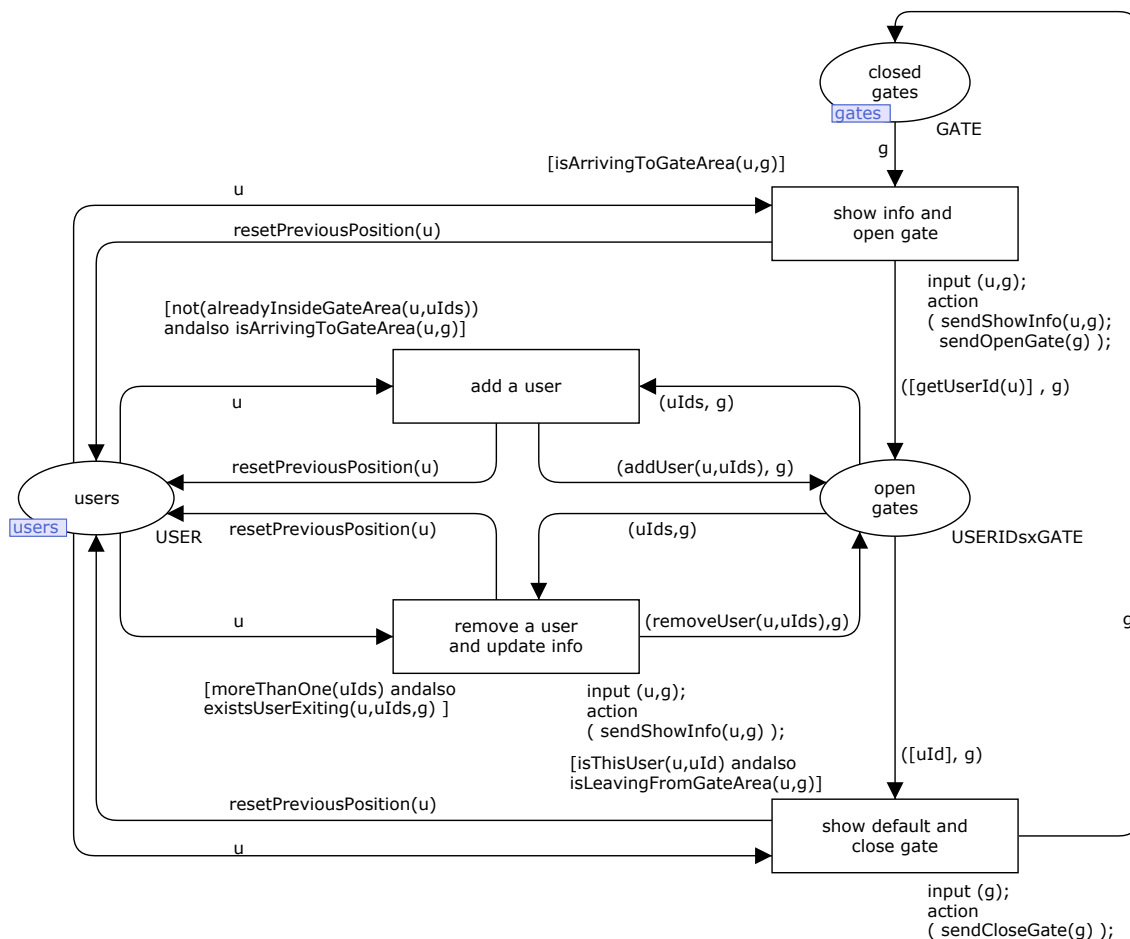


Figura 4. Módulo CPN para o dispositivo *porta de entrada*.

e *close* que são recebidas.

Uma vez o servidor OpenSimulator e a CPN Tools instalados e configurados, o *componente de comunicação/execução* e o modelo CPN carregados, passa a ser possível a exploração do mundo virtual.

Existem dois modos possíveis de utilização do APEX. O modo de desenvolvimento e o modo de utilizador. No primeiro modo, o *developer* tem de alterar/estender o modelo CPN base de modo a reflectir o comportamento de novos dispositivos presentes no ambiente. No caso da definição de novos dispositivos, pequenas modificações tem também de ser efectuadas no módulo de comunicação/execução (módulo C#) de modo a que este saiba como reflectir no novo objecto do ambiente as instruções do modelo CPN. No modo de utilizador, este somente tem de descarregar um visualizador e configurá-lo de modo a poder aceder através da Web ao ambiente simulado. O acesso ao servidor é efectuado através da utilização de uma conta de utilizador criada para o efeito. Nesse modo são utilizados modelos CPN e ambientes virtuais já desenvolvidos sendo que o utilizador experimenta o sistema através de uma simples conexão ao servidor.

Para além disso, a criação, remoção e modificação de objectos pode ser efectuada através da ferramenta de

visualização (ver figura 6) *gritando* (forma de interacção do avatar com o mundo disponibilizada pelo visualizador) os seguintes comandos:

- load-oar ficheiro - responsável por carregar um ambiente/objecto a partir de um ficheiro;
- save-oar ficheiro - responsável por guardar um ambiente/objecto num ficheiro;
- clear - remove todos os objectos presentes no ambiente, ficando somente os utilizadores e o terreno.

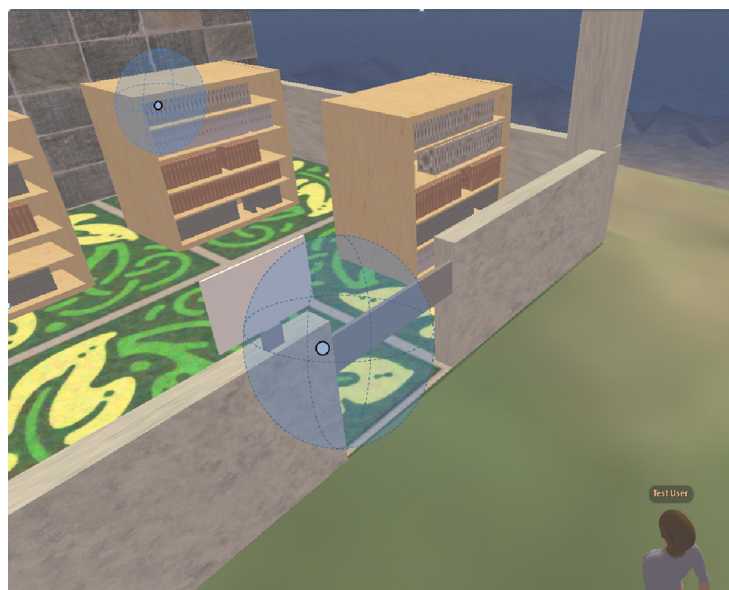
## 6. Conclusões e trabalho futuro

A experiência dos utilizadores em ambientes ubíquos é um factor determinante no seu sucesso. Possibilitar a exploração precoce das características de tais sistemas irá ajudar a antecipar potenciais problemas de utilização e reduzir o custo de *re-design*. No entanto, o desenvolvimento de protótipos no ambiente físico de destino é na maior parte dos casos inviável. Isto acontece devido aos custos de desenvolvimento de tais protótipos e por poder levar a uma ruptura do ambiente a que o sistema em desenvolvimento se destina. Alternativas devem ser exploradas, fornecendo um adequado grau de experiência no uso de sistemas ubíquos, mas evitando os custos da implementação real.

```
foreach (KeyValuePair<Scene, List<SceneObjectGroup>> kvp in HelloWorldModule.scene_prims)
{
    foreach (SceneObjectGroup sog in kvp.Value)
    {
        //Object: gate
        if (sog.Name.Equals("gate"))
        {
            if (action.Equals("open")) //Open the gate
                sog.AbsolutePosition = new Vector3(130.7f,130.9f,25.9f)
            else //Close the gate
                if (action.Equals("close"))
                    sog.AbsolutePosition = new Vector3(128.6f,130.9f,25.9f)
        }

        //Object: screen
        ...
    }
}
```

**Figura 5. Código relativo ao comportamento dos objectos do OpenSimulator**



**Figura 6. Interface do visualizador**

Este artigo apresenta uma proposta para tais alternativas. Uma plataforma de prototipagem baseada em simulação para sistemas de computação ubíquos. Esta plataforma fornece, juntamente, a expressividade das redes de Petri, com a possibilidade de exploração de uma simulação virtual 3D do sistema modelado. O desenvolvimento de modelos e ambientes 3D é acelerado através da utilização do modelo CPN base e de ambientes pré-definidos. Isto permite que potenciais utilizadores possam explorar a simulação do sistema antes da sua implementação, sendo possível ter uma abordagem de prototipagem com um baixo custo.

Desenvolvimentos futuros da plataforma envolvem a sua avaliação com os utilizadores e *developers*. A avaliação dos utilizadores refere-se à fidelidade dos resultados, i.e. se os ambientes prototipados podem ser usados eficazmente permitindo aos utilizadores uma experiência suficientemente rica do *design*. A avaliação dos *developers*

foca-se na agilidade da abordagem, na facilidade com que os protótipos podem ser desenvolvidos de uma forma precisa para ambientes ubíquos.

Trabalhos futuros no desenvolvimento da plataforma referem-se a questões técnicas que visam melhorar o suporte a utilizadores e *developers*. Estes incluem:

- A possibilidade de adicionar novos utilizadores à simulação em tempo de execução, sem que estes tenham que estar previamente inseridos no modelo;
- Reduzir a quantidade de informação trocada entre a CPN Tools e APEX para a mínima possível – isto é relevante para prever que a CPN Tools execute além dos recursos;
- Conectar a simulação a dispositivos móveis, tais como PDAs, via bluetooth – isto irá permitir uma ex-

perícia de utilização mais realista e imersiva.

## Agradecimentos

Este trabalho é suportado pela Fundação para a Ciência e Tecnologia (FCT, Portugal) através da bolsa de doutoramento SFRH/BD/41179/2007.

## Referências

- [Abowd 05] G.D. Abowd, G.R. Hayes, G. Iachello, J.A. Kientz, S.N. Patel, M.M. Stevens, e K.N. Truong. Prototypes and paratypes: designing mobile and ubiquitous computing applications. *IEEE Pervasive Computing*, 4(4):67–73, 2005.
- [Braubach 02] L. Braubach, A. Pokahr, D. Moldt, A. Bartelt, e W. Lamersdorf. Tool-supported interpreter-based user interface architecture for ubiquitous computing. Em *Interactive Systems*, volume 2545 de *Lecture Notes in Computer Science*, páginas 89–103. Springer-Verlag, 2002.
- [Buchenau 00] M. Buchenau e J.F. Suri. Experience prototyping. Em *Proceedings Designing Interactive Systems (DIS'00)*, páginas 424–433. ACM Press, 2000.
- [Davies 05] N. Davies, J. Landay, S. Hudson, e A. Schmidt. Rapid prototyping for ubiquitous computing — guest editors' introduction. *IEEE Pervasive Computing*, 4(4):15–17, 2005.
- [Disz 97] T.E. Disz, M.E. Papka, e R. Stevens. UbiWorld: an environment integrating virtual reality, supercomputing, and design. Em *Proceedings of the Heterogeneous Computing Workshop*, páginas 46–59, April 1997.
- [Gallasch 01] G. Gallasch e L.M. Kristensen. Comms/CPN: A communication infrastructure for external communication with design/CPN. Em K. Jensen, editor, *3rd Workshop and Tutorial on Practical Use of Coloured Petri Nets and the CPN Tools (CPN'01)*, páginas 75–90. DAIMI PB-554, Aarhus University, 2001.
- [Garlan 02] D. Garlan, D.P. Siewiorek, A. Smailagic, e P. Steenkiste. Project Aura: toward distraction-free pervasive computing. *IEEE Pervasive Computing*, páginas 22–31, April-June 2002.
- [Harrison 08] M.D. Harrison, J.C. Campos, G. Doherty, e K. Loer. Connecting rigorous system analysis to experience centred design. Em E. Law, E. Hvannberg, e G. Cockton, editores, *Maturing Usability: Quality in Software, Interaction and Value*, Human Computer Interaction Series, páginas 56–74. Springer-Verlag, 2008.
- [Harter 01] A. Harter, A. Hopper, P. Steggles, A. Ward, e P. Webster. The anatomy of a context-aware application. *Wireless Networks*, 1:1–16, 2001.
- [Jensen 07] K. Jensen, L.M. Kristensen, e L. Wells. Coloured petri nets and CPN tools for modelling and validation of concurrent systems. *International Journal on Software Tools for Technology Transfer (STTT)*, 9(3-4):213–254, 2007.
- [Li 04] Y. Li, J.I. Hong, e J.A. Landay. Topiary: a tool for prototyping location-enhanced applications. Em *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, páginas 217–226. ACM, 2004.
- [Massink 99] M. Massink, D.J. Duke, e S.P. Smith. Towards hybrid interface specification for virtual environments. Em D.J. Duke e A. Puerta, editores, *Design, Specification and Verification of Interactive Systems '99*, páginas 30–51. Springer-Verlag, 1999.
- [Navarre 05] D. Navarre, P. Palanque, R. Bastide, A. Schyn, M. Winckler, L. Nedel, e C. Freitas. A formal description of multimodal interaction techniques for immersive virtual reality applications. Em *INTERACT 2005*, volume 3585 de *Lecture Notes in Computer Science*, páginas 170–183. Springer-Verlag, 2005.
- [O'Neill 09] E. O'Neill, D. Lewis, e O. Conlan. A simulation-based approach to highly iterative prototyping of ubiquitous computing systems. Em *2nd International Conference on Simulation Tools and Techniques*, páginas 1–10. ICST, 2009.
- [Shirehjini 05] Ali A. Nazari Shirehjini e Felix Klar. 3DSim: rapid prototyping ambient intelligence. Em *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*, 2005.
- [Smith 99] S. Smith, D. Duke, e M. Massink. The hybrid world of virtual environments. *Computer Graphics Forum*, 18(3):C287–C307, 1999.

# Locomoção Virtual via Movimento Físico do Utilizador: Controlo de Velocidade do Avatar

Luís Bruno  
INESC-ID, ESTIG  
Av. Rovisco Pais, 1000 Lisboa  
lbruno@ipbeja.pt

Bruno de Araújo  
INESC-ID  
Av. Rovisco Pais, 1000 Lisboa  
brar@vimmi.inesc-id.pt

João Pereira  
INESC-ID  
Av. Rovisco Pais, 1000 Lisboa  
jap@inesc-id.pt

Joaquim Jorge  
INESC-ID  
Av. Rovisco Pais, 1000 Lisboa  
jaj@inesc-id.pt

## Sumário

*A navegação é uma das tarefas principais que os utilizadores podem realizar nos ambientes virtuais 3D. Muitas técnicas, metáforas e dispositivos de interacção têm sido desenhados, implementados e avaliados nesta área da Interacção Pessoa/Computador. Inicialmente, foram usados métodos de interacção indirecta para controlar a navegação, baseados no uso de controlos virtuais e físicos, mas, actualmente, os métodos inovadores que têm sido mais comumente investigados são suportados em interacção directa, conduzida pela acção do controlo de partes do corpo do utilizador, como a cabeça, o tronco, os braços/mãos e as pernas/pés. Este artigo descreve uma abordagem de uma técnica de navegação virtual denominada “Caminhar no Lugar” que permite a um utilizador controlar a velocidade de deslocamento do avatar. Fizem-se experiências para determinar uma forma de cálculo genérica da velocidade instantânea da locomoção horizontal virtual baseada na velocidade vertical dos pés do utilizador ao caminhar no lugar, através do correlacionamento com a velocidade horizontal do utilizador no caminhar real. Os resultados obtidos permitiram determinar uma fórmula do cálculo da velocidade horizontal virtual para cada tipo de passo (devagar, moderado e rápido) e uma outra mais genérica e independente do tipo de passo.*

## Palavras-chave

*Interfaces 3D com o utilizador, caminhar no lugar, locomoção virtual e ambientes virtuais*

## 1. INTRODUÇÃO

A navegação em Ambientes Virtuais (AV's) tem sido objecto de investigação por parte da comunidade científica da área, através da criação e da avaliação de novas técnicas e metáforas de interacção que permitam aos utilizadores movimentarem o seu observador virtual (avatar). A navegação 3D pode ser definida como o processo de movimentação do observador virtual entre duas posições consecutivas num AV, resultantes das acções de controlo do utilizador. Estas podem ser executadas utilizando métodos indirectos, que fazem uso de dispositivos físicos de entrada de dados (ex: *joysticks*) ou de controlos virtuais (*widgets*), ou usando métodos directos, que são suportados nos movimentos dos segmentos corporais do utilizador. O caminhar real [Usoh99] [Steinicke08] é o processo mais natural e realista para os utilizadores controlarem a sua locomoção virtual num AV. É uma técnica que não requer aprendizagem, não desfoca os utilizadores dos seus objectivos principais na realização das tarefas e permite mais naturalmente construir um modelo mental da sua

localização no ambiente. No entanto, é virtualmente impossível aos sistemas de *tracking* (seguimento) seguirem os movimentos dos utilizadores em áreas e volumes de dimensões ilimitadas. Assim, devido a estas limitações foi desenvolvida uma outra técnica de controlo da locomoção virtual denominada “caminhar no lugar” (WIP – Walking-in-place) [Usoh99][Templemen99] [Razaque02] [Yan04] [Templemen06] [Feasel08] [Wendt10], que também utiliza os movimentos corporais dos pés/pernas, mas que restringe o alcance da movimentação do utilizador a uma área limitada de acção. É uma técnica que permite, à semelhança do caminhar real, um tipo de percepção designada por propriocepção, que recebe a informação dos músculos e dos tendões aquando dos seus movimentos, o que permite às pessoas saberem o posicionamento de partes do seu corpo ou a localização relativa do corpo como um todo em relação ao ambiente envolvente.

Esta técnica tem sido objecto de desenvolvimento de vários sistemas nos últimos anos, com destaque para o “Gaiter” [Templemen99], o “LLCM-WIP” [Feasel08] e o

“GUD WIP” [Wendt10]. Estes sistemas podem ser utilizados em conjunto com diferentes sistemas de visualização: HMD (Head Mounted Display), CAVE [Razaque02] e Wall (ecrã de grandes dimensões). Os principais problemas identificados nos sistemas que utilizam a técnica “caminhar no lugar” estão relacionados com a correcta detecção do início e do fim da marcha, com a redução dos tempos de latência do início e do fim da navegação virtual, e com a suavidade do movimento virtual produzido (e respectiva representação visual) entre passos consecutivos e no decorrer de cada um deles.

As duas acções principais desenvolvidas pelas pessoas durante o caminhar são o controlo da direcção e da velocidade do seu movimento [Templeman99]. Poucos trabalhos têm sido desenvolvidos [Bouguila02] [Bouguila04] relacionados com o controlo da direcção da locomoção virtual, baseada somente nas posições/orientações dos pés/pernas dos utilizadores. Normalmente, estes sistemas socorrem-se do uso de dispositivos físicos de controlo (ex: *joysticks*) para fornecerem a direcção da locomoção ou alternativamente usam a orientação de outros segmentos corporais (torso, peito, cabeça e mão) para esse efeito.

Utilizando o movimento vertical dos pés ao caminhar no lugar, os utilizadores esperam que o sistema seja sensível às variações da velocidade vertical desse movimento, por forma a assegurar um maior realismo no controlo da velocidade de locomoção do avatar, designada daqui em diante por velocidade virtual. O trabalho de [Feasel08] ajusta a velocidade vertical dos pés à velocidade virtual, através da identificação de um factor escala para cada um dos utilizadores do sistema. Para tal, é medida a velocidade média do utilizador num teste de caminhar real em passo de cadência moderada e feito um teste de caminhar no lugar com a mesma cadência. Da relação entre as duas velocidades é obtido o factor escala que é utilizado na obtenção da velocidade virtual.

O trabalho de investigação deste artigo tem como objectivo obter uma fórmula genérica que permita determinar um valor realista da velocidade virtual, suportada na velocidade vertical dos pés, e que contemple diferentes tipos de passos (lentos, moderados e rápidos). Esse valor pode ser importante para incrementar o grau de imersão e de qualidade da experiência de locomoção em ambientes virtuais por parte dos utilizadores. Para tal, parte-se da hipótese de que se as velocidades verticais dos pés do caminhar no lugar e do caminhar real forem semelhantes, então será a relação entre a velocidade horizontal e a velocidade vertical dos pés que determinará a componente fixa dessa fórmula.

Para testar esta hipótese e para determinar a fórmula, foi desenvolvida uma experiência em que foram recolhidos dados de seis participantes, realizando várias sessões de caminhar no lugar e em frente, e utilizando três tipos de passos (lentos, moderados e rápidos). Nestas sessões, foi utilizado um reconhecedor de “caminhar no lugar”, baseado numa máquina de estados, que foi também adaptada como reconhecedor do caminhar real.

Associado a este sistema, foi desenvolvido um módulo de análise e produção dos dados estatísticos que suportam os resultados deste trabalho.

Este artigo descreve no capítulo 2 o trabalho relacionado e o estado da arte desta linha de investigação e da área da biomecânica. O capítulo 3 apresenta a arquitectura do sistema reconhecedor do caminhar no lugar e os seus respectivos passos de implementação. No capítulo 4 são apresentados o guião e as condições em que a experiência do registo de dados ocorreu. No capítulo 5 são apresentados os resultados estatísticos da experiência e a respectiva discussão para os elementos mais relevantes. O capítulo 6 apresenta as conclusões do trabalho, identificando os resultados positivos obtidos e algumas das suas limitações. Neste capítulo, são propostos vários trabalhos futuros que podem melhorar algumas das lacunas identificadas.

## 2. TRABALHO RELACIONADO

A locomoção, que é um caso particular da navegação em espaços físicos, é uma característica dos animais que lhes permite deslocarem-se de um ponto para outro. No caso dos seres humanos, a locomoção não se cinge ao caminhar, podendo integrar outros tipos de deslocação como são o correr, o saltar ou o sapatear, entre outros. As duas acções principais desenvolvidas pelas pessoas durante o caminhar são o controlo da direcção e da velocidade do seu movimento [Templeman99]. Para controlar a locomoção do observador virtual têm sido investigadas diferentes técnicas, baseadas em interacção indirecta (controlos físicos e virtuais) e interacção directa em que se utilizam os movimentos corporais de partes do seu corpo para controlar a tarefa. Nesta linha de investigação, foram apresentadas diferentes metáforas do controlo da locomoção virtual, baseadas no uso das mãos (acção de apontar) [Bowman97], do torso, do olhar [Bowman97], do caminhar no lugar (*WIP interfaces*) [Usuh99] [Templeman99] [Feasel08] e do caminhar real [Usuh99] [Steinicke08]. Ainda no suporte a interfaces de locomoção, foram também desenvolvidos vários sistemas mecânicos de suporte à locomoção física e virtual, como são o caso das passadeiras rolantes (*treadmill*) lineares [Slater93] e omni-direccionais [Darken97], de dispositivos com pedais (ex: bicicletas), cadeiras de rodas, *skates*, tapetes com sensores, e outros dispositivos com sensores de pressão que detectam a pressão e a posição dos pés dos utilizadores [Bouguila02] [Bouguila04].

No respeitante às interfaces WIP que são objecto de interesse desta investigação, muitas abordagens têm sido desenvolvidas:

[Slater95] apresentou uma técnica interactiva para a navegação num AV suportada numa rede neuronal, que detecta os padrões dos movimentos da cabeça quando os utilizadores estão a caminhar no lugar a dar passos. Quando o sistema detecta um passo, então é iniciado o movimento virtual do avatar na direcção do seu olhar.



[Templemen99] desenvolveu um sistema chamado "Gaiter", que permite aos utilizadores controlarem o seu movimento num AV de grandes dimensões, com base na análise da extensão da componente horizontal do movimento do joelho.

[Razaque02] desenvolveu técnicas de redireccionamento da posição e da orientação do utilizador, quando ele está a caminhar no lugar, no espaço de *tracking* de uma CAVE de três paredes. O objectivo do uso destas técnicas é impedir que o utilizador olhe para a parede em falta, assegurando-lhe assim um maior sentido de imersão.

[Yan04] descreve um sistema de caminhar no lugar, que calcula a velocidade virtual baseada na velocidade vertical dos pés, reduzindo o tempo de latência do sistema comparativamente com outros sistemas anteriores.

[Feasel08] apresenta um sistema denominado "LLCM-WIP", que pretende reduzir a latência do início e do fim do caminhar no lugar, a suavidade da locomoção entre passos e o controlo contínuo da velocidade durante cada passo. Este sistema faz uso da orientação do peito para controlar a direcção da locomoção.

[Wendt10] descreve um sistema de caminhar no lugar, denominado "GUD WIP", que gera a velocidade de locomoção virtual baseada no ciclo biomecânico do caminhar real. Os resultados do controlo da velocidade virtual deste sistema foram comparados com os do caminhar real e com o sistema "LLCM-WIP" [Feasel08]. Essa avaliação demonstrou que a velocidade gerada pelo "GUD WIP" tem uma consistência muito maior com o caminhar real do que o "LLCM-WIP" e que existe uma consistência grande entre a frequência dos passos e a velocidade tanto no caminhar real como no "GUD WIP".

Em seguida, são descritos alguns conceitos da área da biomecânica, que ajudam a contextualizar este trabalho desenvolvido. Um ciclo completo de marcha durante o caminhar real é definido pelo período que decorre, desde o contacto de um pé no solo até ao contacto seguinte desse mesmo pé. Esse ciclo é composto por duas fases principais para cada pé [Sousa08]: (i) fase de apoio (*Stance*) e (ii) fase oscilatória (*Swing*), conforme se pode ver na Fig. 1.

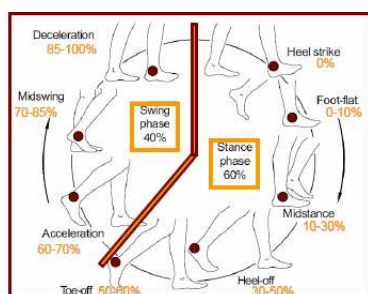


Fig. 1 – Fases do ciclo de marcha ao caminhar

Por sua vez, na fase oscilatória, cada pé pode assumir diferentes estados: (i) Elevação que corresponde a um período de aceleração e (ii) Descida no qual ocorre um período de desaceleração.

### 3. DESCRIÇÃO DO SISTEMA RECONHECEDOR DO CAMINHAR NO LUGAR

O reconhecedor que suporta esta investigação recebe os dados de um sistema de *tracking* (seguimento), baseado na tecnologia de infra-vermelhos, que segue a posição e orientação dos dois pés do utilizador. A ponta de cada pé foi a zona escolhida para detectar o movimento vertical dos pés e consequentemente para calcular a sua velocidade vertical. Essa zona dos pés é das que têm menos oclusões para um sistema de *tracking* como o utilizado neste trabalho. Em cada ponta do pé, é colocado um corpo rígido com quatro marcadores. O reconhecedor realiza várias operações de processamento de sinal sobre as amostras de dados da posição e da orientação de cada pé para calcular a velocidade virtual, conforme descrição mais detalhada neste capítulo.

No processo de reconhecimento do caminhar são identificados os diferentes estados que cada pé pode assumir. Para tal, foi implementada uma máquina de estados, à semelhança de [Wendt10], que reflecte de forma genérica o ciclo da marcha humana. Cada pé tem um estado que pode ser de apoio ou oscilatório. A marcha tem também um estado geral dependente dos estados dos dois pés em conjunto. Se os dois pés estão em apoio (duplo apoio), então o estado da marcha é "Parado". Se um dos pés está em oscilação e o outro está em apoio, então o estado da marcha é "Caminhando". O Diagrama de Transição de Estados, visível na Fig.2, descreve esses estados e as respectivas condições de activação.

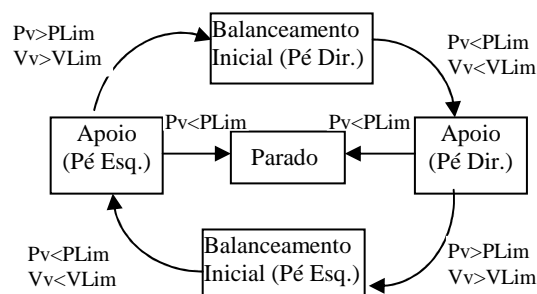
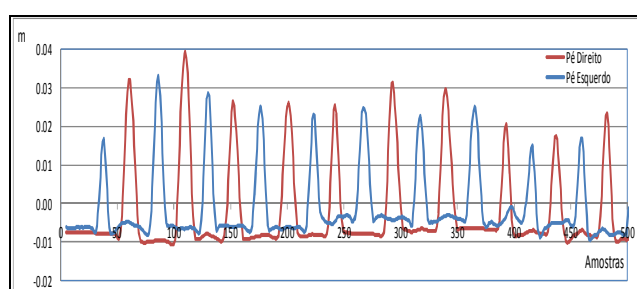


Fig. 2 – Diagrama de transição da máquina de estados

Para detectar o estado de oscilação dos pés foram utilizadas duas condições: (i) a posição vertical (**Pv**) de cada pé tem de ser superior a um determinado limiar mínimo de altura (**PLim**) e (ii) a velocidade vertical (**Vv**) de cada pé tem de ser superior a um limiar mínimo de velocidade (**VLim**). Para identificar a passagem da fase de oscilação para a fase de apoio, as condições são as inversas das enunciadas previamente. O utilizador está no estado parado quando os dois pés estão em duplo apoio e a condição para activar esse estado é quando a altura dos dois pés está abaixo do limiar mínimo **PLim**. Com base na análise dos dados obtidos, foi identificado o valor de 0.06 m/s como sendo a velocidade **VLim**. Já o limiar mínimo de altura do pé (**PLim**) terá de ser identificado caso a caso, para cada utilizador e em cada sessão, visto que o corpo rígido colocado na ponta de cada pé tem

alturas diferentes, conforme os pés e os sapatos dos utilizadores.

Os dados recebidos pelo sistema de *tracking* têm uma taxa de amostragem da ordem dos 60hz, e embora o sinal produzido não apresente taxas de frequências elevadas, foi necessário aplicar um filtro aos dados recebidos. Assim, foram experimentados dois tipos de filtros: médias dos  $n$  últimos valores (3, 5 e 10 últimas amostras) recebidos previamente e o filtro Butterworth(BW) passa-baixo de 4ª ordem (usado em [Feasel08]) e com uma frequência de corte de 5hz. Da análise de dados efectuada, verificou-se que a média das cinco últimas amostras (ver o gráfico da **Fig.3**) produzia uma suavização semelhante à do filtro BW, tinha menos custo computacional e o tempo de latência de início e fim do movimento era inferior.



**Fig. 3** – Gráfico do sinal da posição vertical dos pés direito e esquerdo aplicado um filtro com a média das últimas 5 amostras

Na **Fig. 3** pode-se ver o sinal produzido por cada um dos dois pés. A onda azul representa os dados do pé esquerdo e a encarnada os do pé direito. O eixo das ordenadas representa a altura que cada um dos pés atinge. Verifica-se alguma irregularidade na altura dos pés na fase de apoio que é normalizada pelo uso do “**PLim**” na máquina de estados.

As operações de reconhecimento do caminhar no lugar e as operações subsequentes do cálculo da velocidade virtual são implementadas seguindo a sequência de passos descrita em baixo:

1. Capturar a posição e a orientação de cada pé;
2. Filtrar os dados com o filtro escolhido;
3. Calcular a velocidade de cada um dos pés, por diferenciação entre duas amostras consecutivas de posição em relação ao tempo, e obter o seu valor absoluto;
4. Calcular a velocidade do corpo do utilizador com base na soma das velocidades verticais dos dois pés;
5. Calcular a direcção do deslocamento;
6. Detectar o estado da marcha com base nas condições da máquina de estados;
7. Escalar a velocidade obtida em 4. com um factor que projecta a velocidade horizontal virtual;
8. Determinar o deslocamento da posição do avatar com base na velocidade calculada em 7.;

9. Integrar a nova posição e o vector de direcção do deslocamento no sistema do avatar;

10. Efectuar o *render* do ponto de vista do avatar.

Para calcular os dados da velocidade do caminhar real em frente, foram feitas algumas alterações no sistema. Assim, o cálculo da velocidade horizontal do corpo é considerado como sendo metade da soma da velocidade horizontal dos dois pés (partindo do princípio de que o corpo está numa posição equidistante em relação aos dois pés). A máquina de estados, neste caso, só tem uma condição para activar e terminar o movimento de cada pé, que é baseada no limiar de velocidade mínima horizontal (**VHLim**). Com base na análise de dados, foi identificado o valor de 1.0 m/s como sendo o valor adequado para o valor de **VHLim**.

#### 4. DESCRIÇÃO DA EXPERIÊNCIA

As sessões de captura de dados da experiência foram realizadas num laboratório que tem o sistema de tracking Optitrack<sub>tm</sub>, baseado em dez câmaras de infra-vermelhos que enviam a informação para o sistema de software TrackingTools<sub>tm</sub><sup>1</sup>. Este sistema é responsável pela recepção e processamento das imagens das várias câmaras e pelo cálculo das posições e orientações dos marcadores e corpos rígidos visíveis pelo sistema. Este sistema envia os dados calculados por rede para as aplicações que os solicitam.

A área de tracking disponível para a realização destes testes é de 3mX3m (conforme se pode ver na **Fig.4**) e tem no seu centro geométrico o referencial do sistema.



**Fig. 4** – Área de *tracking* do laboratório onde decorreu a experiência

A aplicação que permitiu fazer o registo, processamento e produção dos resultados estatísticos foi baseada na plataforma OpenIVI<sup>2</sup> (projecto de código aberto)[Araújo05], que permite desenvolver, de uma forma rápida, aplicações que usem técnicas de visualização avançada 3D, combinadas com técnicas de interacção inovadora. Este sistema é constituído por vários módulos que lhe permitem funcionar em ambiente distribuído e multi-threading. Os módulos de gestão de eventos e de eventos OT (baseado no projecto de código

<sup>1</sup>URL do OptiTrack e TrackingTools:  
<http://www.naturalpoint.com/optitrack/>

<sup>2</sup> URL do OPENIVI: <http://open5.sourceforge.net/>

aberto “OpenTracker<sup>3</sup>”) são os responsáveis pela distribuição das mensagens de dados que o sistema recebe do sistema de *tracking*. O módulo de interacção é o responsável pela interpretação dos eventos de entrada e pelo processamento das respostas e foi utilizado neste caso para produzir os dados estatísticos que suportam os resultados da experiência. O módulo de visualização suportado na tecnologia de código aberto “OpenSG”<sup>4</sup> permitiu visualizar a locomoção virtual resultante dos dados de entrada capturados.

O equipamento utilizado onde foi executada a aplicação anterior foi um computador com o S.O. Microsoft WindowsXP, processador Intel Core2 2.80GHz, placa gráfica NVIDIA GeForce 9600M GT e com 4GB de memória.

As sessões de obtenção de dados da experiência contaram com a presença de seis participantes (cinco do género masculino e um do género feminino), os quais foram instruídos para realizar um conjunto de nove sessões de caminhar real em frente e outro conjunto com nove sessões de caminhar no lugar. Cada conjunto de nove sessões teve três tipos diferentes de passos: três lentos, três moderados e três rápidos. Estas experiências tiveram sequências alternadas dos tipos de passos, para que não houvesse enviesamento dos resultados. Em cada pé dos participantes foi colocado um corpo rígido constituído por quatro marcadores reflectores, conforme se pode ver na **Fig.5**, e que permitiu ao sistema de *tracking* registar a sua posição e a orientação em relação ao referencial do sistema de *tracking* da sala.



**Fig. 5** – Colocação dos corpos rígidos nos pés

A experiência iniciou-se com a colocação dos marcadores nos pés dos utilizadores e foi-lhes explicada a sequência de tarefas que eles teriam de realizar. Não lhes foram fornecidos exemplos dos diferentes tipos de passos que se pretendiam registar, dando aos participantes liberdade para escolherem o que eles consideram passos lentos, moderados ou rápidos. No entanto, foi-lhes sugerido que seguissem uma velocidade semelhante para um mesmo tipo de passo nas duas condições: caminhar no lugar e caminhar real. Para capturar os dados do caminhar real em frente, foi pedido aos participantes que se colocassem

atrás do limite da área de *tracking* e que após as instruções do supervisor da experiência iniciassem a sua marcha em frente com o tipo de passo pedido. O supervisor terminou o registo de cada tentativa, após o participante ter ultrapassado o outro limite da zona de *tracking* visível. Verificou-se que os utilizadores, em média, e com base nas limitações do sistema de *tracking*, deram dois a três passos rápidos, quatro a cinco passos moderados e seis a sete passos lentos. Após a realização das nove sessões de caminhar em frente em sequências alternadas, o participante inicia as nove sessões de caminhar no lugar. Ele é posicionado no centro da área de *tracking* e são-lhe dadas as instruções sobre as tarefas a realizar em seguida. O participante inicia cada uma das sessões após a ordem do supervisor da experiência, realizando dez passos no lugar, do tipo previamente solicitado.

Os dados do *tracking* resultantes de cada sessão ficam registados num ficheiro em formato texto plano, com formatação lógica CSV (*Comma Separated Value*) em que cada linha tem os seguintes campos de dados: *timestamp* da amostra, posição( $x_e, y_e, z_e$ ) do pé esquerdo, orientação( $x_e, y_e, z_e, w_e$ ) do pé esquerdo, posição( $x_d, y_d, z_d$ ) do pé direito e orientação( $x_d, y_d, z_d, w_d$ ) do pé direito.

## 5. RESULTADOS E DISCUSSÃO

Os dados resultantes das várias sessões da experiência foram processados e validados para obter os resultados estatísticos presentes neste capítulo. Os dados do caminhar no lugar foram considerados consistentes e utilizados na sua totalidade. Nem todos os dados registados do caminhar real em frente foram utilizados na produção dos resultados. Como neste tipo de caminhar, os utilizadores iniciavam as sessões atrás da zona de *tracking*, alguns registos das amostras iniciais não tinham valores consistentes, porque os marcadores estavam temporariamente não visíveis para as câmaras de infravermelhos. Para tal, foram removidas algumas amostras que poderiam desvirtuar a consistência dos resultados finais.

A **Fig.6** mostra os resultados totais dos dados das sessões do caminhar real em frente. Em baixo são discutidos e justificados os resultados evidenciados.

A velocidade do corpo do utilizador ao longo do plano horizontal do chão (**VelZF**) varia incrementalmente (como era expectável), desde os passos lentos (1.08 m/s), passando pelos moderados (1.44 m/s) e até aos rápidos (2.31 m/s). No entanto, o desvio padrão dos passos lentos permite inferir que houve uma maior variabilidade na velocidade dos passos lentos, talvez devido ao facto de os utilizadores terem adoptado diferentes interpretações e comportamentos sobre o conceito de passos lentos. Verifica-se também uma diferença muito maior na relação entre velocidades dos passos rápidos e moderados (1.60), do que na relação entre estes últimos e os lentos (1.34), conforme se pode ver na **Fig.7**. Segundo [Dean65] a velocidade do caminhar real varia exponencialmente com a frequência dos passos. Assim, é provável que os

<sup>3</sup> URL do OpenTracker: <http://studierstube.icg.tu-graz.ac.at/opentracker/>

<sup>4</sup> URL do OpenSG : <http://www.opensg.org/>



utilizadores tenham adoptado uma frequência de passo muito mais elevada do que a do passo moderado, o que levou a um aumento exponencial da velocidade.

	VelZF(m/s)		VelYF(m/s)		Rel VelZF/VelYF(m/s)	
	AVG.	STDEV.	AVG.	STDEV.	AVG.	STDEV.
Passos Lentos	1.08	25%	0.33	22%	3.41	11%
Passos Moderados	1.44	20%	0.41	15%	3.51	11%
Passos Rápidos	2.31	14%	0.50	6%	4.70	12%
Todos os passos	1.61	39%	0.41	21%	3.87	19%

**Fig. 6** – Resultados dos dados do caminhar em frente por tipo de passo

A velocidade vertical (**VelYF**) dos pés ao caminhar em frente demonstra que a sua variação é inversamente proporcional ao tipo de passo realizado, conforme se pode ver na **Fig.7**. Assim, a variação da relação dos passos moderados/passos lentos (1.27) é superior à variação da relação dos passos rápidos/passos moderados (1.20), embora os seus valores não sejam muito distintos.

	VelZF	VelZF/VelYF	VelYF	VelYL
Passos Moderados/Passos Lentos	1.34	1.03	1.27	1.35
Passos Rápidos/Passos Moderados	1.60	1.34	1.20	1.22

**Fig. 7** – Relações das variações de velocidades entre diferentes tipos de passos

Do quadro da **Fig.6** é importante salientar que a relação entre a velocidade **VelZF/VelYF** se mantém muito semelhante para os passos lentos (3.41m/s) e moderados (3.51m/s), mas que é muito superior para os passos rápidos (4.70m/s). Essa variação também é evidente na **Fig.7**, em que esse valor cresce de 1.03 (passos moderados/passos lentos) para 1.34 (passos rápidos/passos moderados).

A **Fig.8** mostra os resultados totais dos dados das sessões de “caminhar no lugar”. A velocidade vertical dos pés (**VelYL**) do caminhar no lugar, expressa na **Fig.8**, demonstra que também existe uma coerência na variação incremental dos seus valores (0.24 m/s, 0.32 m/s e 0.39 m/s) na sequência da realização dos três tipos de passos (lento, moderado e rápido). No entanto, verifica-se que existe uma maior variabilidade nos resultados obtidos, expressa pelos maiores valores do desvio padrão em relação às sessões de caminhar em frente. Isto pode dever-se ao facto de os utilizadores não terem uma noção clara dos tipos de passos no lugar a adoptar, que tenham similaridade com os passos do caminhar em frente. Essa falta de mapeamento levou-os, certamente, a terem diferentes comportamentos entre eles e até nas suas próprias tentativas para os mesmos tipos de passos.

	VelYL(m/s)	
	AVG.	STDEV.
Passos Lentos	0.24	45%
Passos Moderados	0.32	51%
Passos Rápidos	0.39	51%
Todos os passos	0.32	24%

**Fig. 8** – Resultados dos dados do caminhar no lugar por tipo de passo

A velocidade vertical dos pés (**VelYL**) ao caminhar no lugar demonstra, à semelhança da **VelYF**, que a sua

variação é inversamente proporcional à rapidez dos passos realizados, conforme se pode ver na **Fig.7**. Assim, o valor da relação dos passos moderados/passos lentos (1.35) é superior ao da relação dos passos rápidos/passos moderados (1.22), embora os seus valores não sejam muito distintos. Nessa figura, verifica-se que existe uma consistência muito grande nos dados, entre as variações das velocidades verticais dos pés no caminhar em frente e no lugar, o que poderá indiciar a obtenção de um valor de referência para a projecção velocidade horizontal virtual a partir da velocidade vertical dos pés.

Após o tratamento dos dados, foi possível identificar uma relação entre a velocidade vertical dos pés (**VelYF**) no caminhar em frente e a respectiva velocidade (**VelYL**) no caminhar no lugar, para cada um dos tipos de passos experimentados. Assim, e com base na **Fig. 9**, é possível verificar que existe uma consistência nos dados obtidos que poderá suportar o cálculo da velocidade horizontal virtual. Embora os valores de **VelYF** sejam em média 1.31 superiores aos de **VelYL**, todos as relações entre os tipos de passos apresentam um valor semelhante (o desvio padrão é de 4%), o que poderá suportar o cálculo da projecção da velocidade horizontal virtual.

		VelYF / VelYL
Passos Lentos		1.37
Passos Moderados		1.29
Passos Rápidos		1.27
Todos os passos	AVG.	1.31
	STDEV.	4%

**Fig. 9** – Relações entre as velocidades verticais do caminhar em frente e no lugar, por tipo de passo.

Com base nos resultados anteriores, e assumindo que a constante **RVY** é a relação entre **VelYF** e **VelYL** (conforme a **Fig. 9**), que a constante **RVF** é a relação entre **VelZF/VelYF** (conforme a **Fig.6**) propõe-se que o cálculo da velocidade horizontal virtual (**VelZP**) projectada para cada tipo de passo seja baseado na seguinte fórmula (1):

$$\text{VelZP} = \text{VelYL} * \text{RVY} * \text{RVF} \quad (1)$$

Calculada a velocidade vertical do caminhar no lugar, é detectado qual o tipo de passo. Assim, o valor da velocidade vertical dos pés (**VelYL**) é multiplicado pela constante **RVY** das velocidades verticais (lugar e frente) obtidas dos testes (**VelYF/VelYL**) (ver **Fig.8**) para cada tipo de passo, e pela relação das velocidades horizontal e vertical do caminhar em frente (**VelZF/VelYF**) (ver **Fig.6**).

Na **Fig.10** são apresentadas as fórmulas de cálculo que estimam a velocidade instantânea de locomoção virtual, referente a uma determinada amostra, durante o caminhar no lugar, e para um determinado tipo de passo. É igualmente apresentada uma fórmula genérica, que pode ser utilizada para qualquer tipo de passo efectuado.

	VelZP (m/s)
Passos Lentos	$VelYL * 1.37 * 3.41$
Passos Moderados	$VelYL * 1.29 * 3.51$
Passos Rápidos	$VelYL * 1.27 * 4.70$
Qualquer Passo	$VelYL * 1.31 * 3.87$

**Fig. 10** – Fórmula de cálculo da velocidade horizontal virtual

No caso de não ser detectado previamente o tipo de passo, a fórmula ( $VelYL * 1.31 * 3.87$ ) utilizada é baseada nos valores médios dos vários tipos de passos obtidos no tratamento de dados da experiência e que permitem determinar a velocidade horizontal virtual independente do tipo de passo.

## 6. CONCLUSÕES E TRABALHO FUTURO

Foi possível demonstrar, através dos dados estatísticos produzidos nesta experiência, que existe uma correlação entre as velocidades verticais dos pés e as velocidades horizontais de locomoção no caminhar em frente e que essa relação poderá ser utilizada para calcular a velocidade horizontal virtual a partir da velocidade vertical dos pés. Concluiu-se também que não existe uma relação próxima de 1 para 1 (ver **Fig.9**) entre a velocidade vertical do caminhar em frente e a do caminhar no lugar, para cada um dos três tipos de passos utilizados (lentos, moderados e rápidos), embora haja uma grande proximidade nessas relações, o que permite utilizar este valor na formação de uma fórmula da velocidade horizontal virtual.

A relação da velocidade horizontal / velocidade vertical (ver **Fig.6**) do caminhar real em frente, relativa aos passos do tipo rápido, apresenta um valor bastante superior relativamente às mesmas relações dos passos lentos e moderados (que são bastante semelhantes entre si). Isto poderá ser justificado pelo facto de a velocidade de locomoção do ser humano crescer exponencialmente, conforme aumenta a frequência dos passos.

Com base nos valores obtidos, não foi possível calcular uma só fórmula que, de forma exacta, e que, independentemente do tipo de passo, fornecesse o valor da velocidade horizontal virtual a partir da velocidade vertical dos pés e de um factor escala. No entanto, foram obtidas fórmulas para cada um dos tipos de passo, baseadas nos valores da velocidade virtual e nas relações  $VelYF/VelYL$  e  $VelZF/VelYF$ . É claro que isto requer a identificação do tipo de passo que está a ocorrer. Este dado pode ser determinado pelo período de tempo que dura uma das fases do ciclo biomecânico (por exemplo, a fase de aceleração durante a oscilação). Tempos mais pequenos significam maior rapidez e vice-versa. É claro, que este processo poderá trazer problemas de latência do sistema na sua resposta às alterações de velocidade provocadas pelas acções do utilizador. Só depois da ocorrência de algumas amostras é que o sistema detecta o tipo de passo que o utilizador está a realizar e só a partir daí é que poderá fazer a compensação desses valores nas amostras seguintes. Foi também calculada uma fórmula genérica do valor da velocidade pretendida, que utiliza os valores médios das relações de velocidade, que poderá

ser uma aproximação ao valor correcto da velocidade de locomoção virtual.

Com este trabalho realizado, existe oportunidade para a realização de vários trabalhos futuros. Para validar estes resultados ou até obter outros mais relevantes, devem de ser utilizados mais participantes no registo de dados desta experiência.

Pretende-se também investigar o cálculo matemático da velocidade horizontal virtual, a partir de uma fórmula genérica que esteja em função de variáveis associadas ao movimento vertical dos pés durante o caminhar no lugar e que reflecta a característica exponencial que essa velocidade tem em função da frequência dos passos efectuados [Dean65].

O objectivo do cálculo destes valores da velocidade virtual é o de conferir ao utilizador, no seu controlo da locomoção, uma noção de realismo muito semelhante à velocidade que ele obtém nas tarefas de caminhar em frente, durante o seu dia-a-dia. Assim, os resultados obtidos nesta experiência deverão ser testados com utilizadores na tarefa de controlo da locomoção num ambiente virtual. Deverão ser realizados testes de usabilidade para verificar se a resposta do sistema com estas fórmulas é realista, e também identificar qual é a noção que eles têm da velocidade e das distâncias percorridas num AV.

## 7. REFERÊNCIAS

- [Dean65] G. A. Dean. An analysis of the energy expenditure in level and grade walking. *Ergonomics*, 8(1):31–47, 1965.
- [Templeman99] J. Templeman, P. Denbrook, and L. Sibert. Virtual locomotion: Walking in place through virtual environments. *ACM Presence* 8, 6, 598–617, 1999.
- [Templeman06] Templeman, J., Sibert, L., Page, R. and Denbrook, P. “Immersive Simulation to Train Urban Infantry Combat,” In *Virtual Media for Military Applications*, 2006.
- [Bowman97] D.A.Bowman, D. Koller, and L.F. Hodges. Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *Proceedings of the Virtual Reality Annual International Symposium*, pages 45-52, 1997.
- [Usoh99] Usoh, M., Arthur, K., Whitton, M., Bastos, R., Steed, A., Slater, M., and Brooks, F. “Walking > Walking-in-Place> Flying, in Virtual Environments,” In *proceedings of ACM SIGGRAPH’99*, 359-364.
- [Steinicke08] F. Steinicke, et al. “A Universal Virtual Locomotion System: Supporting Generic Redirected Walking and Dynamic Passive Haptics within Legacy 3D Graphics Applications,” In *Proceedings of IEEE Virtual Reality*, 2008.
- [Feasel08] J. Feasel, J. D. Wendt, and M. C. Whitton. LLCM-WIP. Low-latency, continuous-motion walk-

- ing-in-place. In Proceedings of the IEEE 3D User Interfaces (3DUI '08), pages 97–104, 2008.
- [Darken97] Darken, R., Cockayne, W., and Carmein D. “The Omni-Directional Treadmill: A Locomotion Device for Virtual Worlds,” Proc. of UIST'97, 213-221, ACM Press, Banff, Canada, October 1997.
- [Slater93] Slater, M., and Usoh, M. “Presence in Immersive Virtual Environments,” In Proceedings of IEEE VRAIS, 90-96, 1993.
- [Slater95] Slater, M., Usoh, M., and Steed, A. “Taking steps: The influence of a walking technique on presence in virtual reality,” ACM Trans. on Human Interaction 2, 3, 201–219, 1995.
- [Bouguila02] Bouguila, L., Ishii, M., and Sato M. “Realizing a New Step-in-place Locomotion interface for Virtual Environment with Large Display System,” in Proceedings of the of the workshop on Virtual Environments, 101-110, 2002
- [Bouguila04] Bouguila, L., Evequoz, F., Courant, M., and Hirsbrunner, B. “Walking-pad: a step-in-place locomotion interface for virtual environments,” Proceedings of the 6th international conference on Multimodal interfaces – ICMI, 77-81, 2004.
- [Razaque02] S. Razzaque, D. Swapp, M. Slater, M. C. Whitton and A. Steed. Redirected Walking in Place. Proceedings of the 8th. Eurographics Workshop on Virtual Environments, 123-130, 2002
- [Yan04] Yan, L., Allison, R. and Rushton, S. “New Simple Virtual Walking Method – Walking on the Spot,” In 8th Annual Immersive Projection Technology (IPT) Symposium Electronic Proceedings, 2004.
- [Wendt10] J. D. Wendt, M. C. Whitton, and F.P. Brooks. GUD WIP: Gait-Understanding-Driven Walking-In-Place. Proceedings of the IEEE Virtual Reality 2010, Pag. 51-58, 2010.
- [Sousa08] A. Sousa. Análise da Marcha Baseada numa Correlação Multifactorial. Tese de Mestrado em Engenharia Biomédica, Universidade do Porto, Faculdade de Engenharia, 2008.
- [Araújo05] B. R. Araújo, T. Guerreiro, R. Jota, J. A. Jorge, J. M. Pereira. LEMe Wall: Desenvolvendo um Sistema de Multi-Projecção. 13º Encontro Português de Computação Gráfica (EPCG), Vila Real, October 2005.

# Rato versus Joystick como dispositivo de entrada para navegação num Ambiente Virtual com um *Head-Mounted Device*

Paulo Dias  
DETI/IEETA  
Univ. de Aveiro  
paulo.diasj@ua.pt

Paulo Santos  
DETI  
Univ. de Aveiro  
a27638@alunos.det.ua.pt

Carlos Ferreira  
DEGEI/CIO  
Univ. de Aveiro  
Univ. de Lisboa  
carlosf@ua.pt

Beatriz Sousa  
Santos  
DETI/IEETA  
Univ. de Aveiro  
bss@ua.pt

---

## Sumário

*Este artigo apresenta uma experiência controlada realizada para comparar dois dispositivos comuns, um rato e um joystick, como dispositivos de entrada durante a navegação num Ambiente Virtual constituído por um labirinto 3D sendo o dispositivo de saída um Head-Mounted Device. Nesta experiência mediram-se o desempenho e a satisfação de 45 participantes, não se tendo observado uma diferença significativa nas duas condições experimentais, o que permite concluir que ambos os dispositivos têm uma usabilidade comparável nas condições da experiência.*

## Palavras-chave

*Ambientes Virtuais, dispositivos de entrada, Head-Mounted Device (HMD), usabilidade*

---

## 1. INTRODUÇÃO

Embora a Realidade Virtual tenha já áreas de aplicação prática, como o entretenimento, o treino ou a cultura, apresenta ainda desafios grandes a nível da Interacção que terão que ser ultrapassados para que possa ser mais divulgada. Assim, é urgente a compilação de directivas que possam ajudar os projectistas a desenvolver novos sistemas que resultem usáveis [Wilson06a, 06b]. Embora, os estudos empíricos com utilizadores sejam difíceis de planear e por em prática, são fundamentais para ajudar a produzir as referidas directivas. Este artigo apresenta um estudo com utilizadores realizado para investigar uma questão que surgiu no decurso de trabalhos anteriores em que se estudou a usabilidade de plataformas de Realidade Virtual de baixo custo [Sousa-Santos09a]. Durante as experiências realizadas nesse âmbito, foram-se registando comentários e sugestões dos utilizadores, sendo recorrente a sugestão de utilização, em tarefas de navegação num ambiente virtual 3D, de um *joystick* em vez de um rato como dispositivo de entrada de uma plataforma que inclui um *Head-Mounted Device* (HMD). Esta sugestão pareceu promissora e decidiu-se averiguar, objectivamente, como se compara o desempenho e a satisfação dos utilizadores, nas condições das experiências anteriores, quando usam um *joystick* e um rato. Neste artigo descreve-se a experiência controlada realizada com esse objectivo e as secções seguintes

apresentam as plataformas e o ambiente virtual usados, o *design* experimental e os principais resultados obtidos. Apresentam-se ainda algumas conclusões.

## 2. PLATAFORMAS E AMBIENTE VIRTUAL

Neste estudo utilizou-se o mesmo Ambiente Virtual em duas plataformas diferentes: um *desktop* comum e uma plataforma que inclui um *Head-Mounted Display* (HMD) com *Head Tracking* como dispositivo de saída (que se designa, a partir de agora, como plataforma HMD).

O *desktop* utilizado era um Pentium 4 a 2.5GHz com uma placa gráfica NVIDIA GeForce FX 5950 Ultra e um monitor *Wide Screen* de 19" com uma resolução de 800x600 *pixels*. A plataforma HMD incluía uns i-glasses SVGA Pro HMD com uma resolução de 800x600 *pixels*, capacidade de estereoscopia, um *field of view* (FOV) de 26°, uma *frame rate* de 60Hz ou 120Hz (correspondentes à sua utilização em mono ou em estéreo respectivamente), um sensor de orientação (*tracker*) InterTrax 2 da InterSense com três graus de liberdade (*yaw*, *pitch* e *roll*) e um PC com uma placa gráfica nVidia Quadro FXGo 1400.

Como ambiente virtual utilizou-se um jogo simples em que os utilizadores tinham que navegar num labirinto para apanhar objectos espalhados ao longo dos corredores. Optou-se por usar um jogo já que este pode incorpo-



rar várias tarefas de navegação que os utilizadores desempenham espontaneamente. Além de se tornar mais atractivo para utilizadores voluntários, um jogo pode ainda ser vantajoso em estudos de usabilidade em Realidade Virtual, pois permite aos utilizadores concentrarem-se nas suas tarefas, não necessitando de instruções frequentes do experimentador. Por outro lado, optou-se por um labirinto, já que outros autores têm usado, com vantagem, labirintos em estudos sobre navegação em Realidade Virtual [Rudle04].

O jogo (ver Figura 1) tinha sido utilizado em estudos anteriores [Sousa-Santos09a] e foi desenvolvido usando o *Open source 3D Graphics Engine OGRE* [OGRE]. Consistia num labirinto (ambiente Maze) com corredores, muito semelhantes, em que os utilizadores tinham que navegar durante 5 minutos, tendo como objectivo recolher 21 objectos. Estes objectos flutuavam ao nível dos olhos e o utilizador apenas necessitava de lhes tocar para os apanhar, recebendo *feedback* visual (desaparecimento do objecto) quando era bem sucedido.

O ambiente Maze utilizado induz os utilizadores a realizar tarefas simples, como andar em frente e dobrar esquinas, bem como tarefas mais complexas como passar através de portas, de acordo com as recomendações de Griffiths *et al.* [Griffiths06]. Usou-se a metáfora *steering* [Bowman04] para a navegação (especificação contínua da direcção do movimento) fazendo o *tracking* da cabeça na plataforma HMD, ou através do rato no *desktop*.



Figura 1. Aspecto do Ambiente Virtual Maze.

### 3. DESIGN DA EXPERIÊNCIA

Esta experiência foi realizada em duas etapas, na primeira participaram 21 estudantes do ensino secundário que frequentaram a Academia de Verão da Universidade de Aveiro em Julho de 2008 e a segunda envolveu 24 estudantes, também do ensino secundário, da Escola Secundária José Estêvão de Aveiro, em Novembro de 2008. A segunda etapa foi realizada para confirmar os resultados obtidos na primeira etapa (brevemente apresentados em [Sousa-Santos09b]). O design experimental foi mantido, servindo a segunda etapa para acumular resultados com mais utilizadores. As principais características da experiência controlada foram as seguintes:

- Hipótese a testar: a mudança do dispositivo de entrada utilizado com o HMD não tem influência no desempenho dos utilizadores.

- Variável independente: dispositivo de entrada utilizado com o HMD (dois níveis - rato ou *joystick*);

- Variáveis dependentes: número de objectos apanhados, velocidade média atingida e número de colisões;

- Variáveis secundárias: ordem de utilização da plataforma, experiência com jogos e dispositivos de entrada;

- Participantes: 45 estudantes do ensino secundário voluntários (29 rapazes, 16 raparigas), organizados em dois grupos: 22 usaram o rato (grupo G1) e 23 usaram o *joystick* (grupo G2)

- Método experimental: *between groups* (i.e. os utilizadores utilizaram apenas um dos dispositivos de entrada, rato ou *joystick*, quando jogaram na plataforma HMD). Como este método tem a desvantagem de poder haver diferenças entre os perfis dos participantes nos dois grupos, e para estabelecer uma referência de desempenho, todos os utilizadores jogaram também no *desktop*.

Esta experiência envolveu várias fases. Na primeira fase, procedeu-se à explicação da experiência aos participantes. Durante a segunda fase, os utilizadores jogaram nas duas plataformas. Na terceira fase, recolheram-se dados quanto ao desempenho, satisfação, conforto, opinião e preferências. A Figura 2 ilustra estas fases, que se descrevem a seguir mais detalhadamente.

1ª fase – Explicação do jogo - informou-se também cada participante que poderia desistir quando quisesse.

2ª fase – Jogo no ambiente virtual Maze - foi usado o jogo nas mesmas condições, variando o dispositivo de entrada usado com o HMD. Nesta experiência os utilizadores jogaram nas duas plataformas, tendo um grupo jogado com o rato e o outro com o *joystick* na plataforma HMD. Como referido, o desempenho dos utilizadores foi medido através do número de objectos apanhados, da velocidade média atingida e do número de colisões com as paredes que foram registados durante o jogo. Com o objectivo de compensar possíveis efeitos de aprendizagem entre as duas vezes que os utilizadores jogam, metade dos utilizadores começou por jogar no *desktop* e outra metade na plataforma HMD. Os utilizadores foram todos observados durante a experiência.

3ª fase – Questionário para estabelecer o perfil dos utilizadores, medir conforto e satisfação e registar preferências e opiniões - a caracterização dos utilizadores foi feita através de um questionário que incluía questões quanto ao género e idade, bem como experiência e frequência de utilização de computadores, jogos e dispositivos de entrada (características que foram consideradas relevantes por poderem influenciar os resultados). Questionou-se também os utilizadores quanto a efeitos adversos (enjoo, tontura e desorientação) e dificuldades (ver e apanhar os objectos, passar através de portas), que possam ter sentido durante o jogo. Colocaram-se também questões relativamente à satisfação e preferências e solicitaram-se sugestões quanto às plataformas e ao jogo.



**Figura 2. Design experimental:** 45 utilizadores jogaram no *desktop* e com o HMD (rato- 23, *joystick*- 22). Foram registadas medidas de desempenho, satisfação e conforto e foi caracterizado o perfil dos utilizadores. Os utilizadores foram observados e deram sugestões.

#### 4. RESULTADOS

Apresentam-se os principais resultados obtidos na experiência. A análise estatística dos dados correspondentes às medidas de desempenho registadas durante o jogo foi efectuada utilizando Análise Exploratória de Dados (EDA) [Hoaglin83] e testes não paramétricos [Conover99] (devido à não normalidade da distribuição das amostras) e realizados com o software Statistica [Statistica].

Para efeito da análise de dados os utilizadores foram organizados em dois grupos:

G1 - grupo que utilizou a plataforma HMD com um rato como dispositivo de entrada;

G2 - grupo que utilizou a plataforma HMD com um *joystick* como dispositivo de entrada.

Ambos os grupos jogaram no *desktop*, pelo que se apresentam os resultados da seguinte forma:

G1\_m - resultados do grupo G1 com o HMD e o rato;

G1\_d - resultados do grupo G1 no *desktop*;

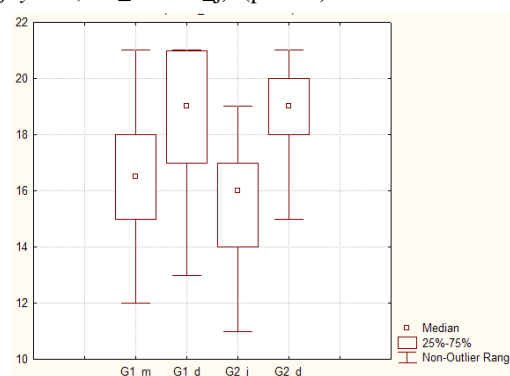
G2\_j - resultados do grupo G2 com o HMD e o *joystick*;

G2\_d - resultados do grupo G2 no *desktop*.

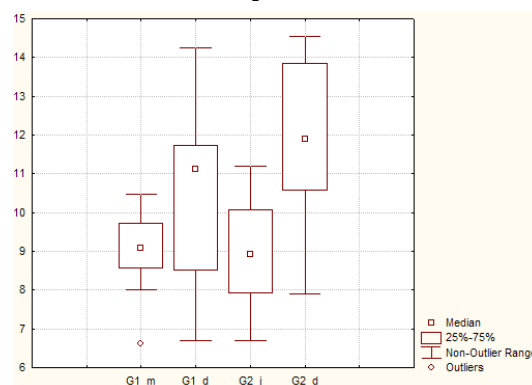
As figuras 3 e 4 mostram as caixas de bigodes correspondentes ao número de objectos apanhados e à velocidade mediana atingida para cada grupo.

Na figura 3 pode observar-se que os utilizadores de ambos os grupos (G1 e G2) apanharam um maior número de objectos quando usaram o *desktop* em relação ao

número apanhado quando usaram o HMD; esta observação foi confirmada pelo teste de Wilcoxon - caso emparelhado - ( $p=0.002$ ,  $p=0.00006$  para o G1 e o G2 respectivamente). Como tínhamos dois grupos de utilizadores diferentes, testou-se também, através do teste de Mann-Whitney, se a diferença entre o desempenho dos dois grupos (G1\_d e G2\_d) no *desktop* era significativa, tendo-se verificado que não era ( $p=0.86$ ). Ou seja, os utilizadores apanharam um número de objectos não significativamente diferente entre os dois grupos quando usaram o *desktop* e apanharam significativamente menos objectos quando usaram a plataforma HMD (com o rato e o *joystick* respectivamente). Também não é significativa a diferença entre o número de objectos apanhados pelos dois grupos quando usaram a plataforma HMD (com o rato e o *joystick*, G1\_m e G1\_j) ( $p=0.18$ ).



**Figura 3. Número de objectos apanhados pelos dois grupos de utilizadores (G1 e G2) ao utilizarem o rato e o *joystick* com o HMD (G1\_m, G2\_j, respectivamente) e o *desktop* (G1\_d e G2\_d).**



**Figura 4. Velocidade média atingida pelos utilizadores dos dois grupos ao utilizarem o *desktop* G1\_d, G2\_d e a plataforma com o HMD (G1\_m e G2\_j).**

A figura 4 mostra os resultados de velocidade média obtidos, que foram análogos aos obtidos para o número de objectos apanhados: os utilizadores atingiram maiores velocidades com o *desktop*, não se observando diferença significativa entre a velocidade atingida com o rato e o *joystick* (G1\_m e G2\_j), nem tão-pouco entre os dois grupos quando usaram o *desktop* (G1\_d e G2\_d); no entanto a diferença de velocidade atingida no *desktop* e

na plataforma HMD (com qualquer dos dois dispositivos) foi considerada significativa pelo mesmo tipo de teste ( $p=0.0036$ ,  $p=0.00004$  para o G1 e o G2 respectivamente).

Estes resultados implicam a aceitação da hipótese nula estabelecida: não se observou diferença significativa entre o desempenho dos utilizadores quando utilizaram a plataforma HMD usando um rato ou um *joystick*.

Quanto à satisfação e preferências demonstradas pelos utilizadores, todos gostaram de jogar em ambas as plataformas; 35 dos 45 preferiram jogar na plataforma HMD, eventualmente devido à novidade, e 10 preferiram o *desktop*, eventualmente devido à experiência que têm com este tipo de plataforma.

Também não se observaram diferenças expressivas nem de preferências nem de satisfação entre os utilizadores que usaram diferentes dispositivos de entrada na plataforma HMD.

## 5. CONCLUSÕES

Apresentou-se uma experiência controlada realizada para averiguar uma questão muito específica referente à usabilidade de uma plataforma com um HMD na execução de um conjunto de tarefas de navegação num ambiente virtual, em relação à qual não foi possível encontrar directivas que permitissem antecipar a resposta. Os resultados obtidos com 45 estudantes do ensino secundário permitiram concluir que, ao contrário do que se podia esperar com base em opiniões de utilizadores em estudos anteriores, não se observou diferença significativa entre o desempenho dos grupos que utilizaram dispositivos de entrada diferentes (rato e o *joystick*). Também quanto a preferências e satisfação, não se verificou diferença expressiva entre os dois grupos.

Apesar das condições particulares da experiência, quanto às tarefas a realizar, e do número não muito elevado de participantes, estamos convencidos que o resultado obtido é generalizável a outras tarefas de navegação em ambientes 3D com um HMD, já que os participantes estão numa faixa etária em que a utilização do *joystick* é muito frequente, o que lhes dá, à partida, vantagem na sua utilização em relação a outros utilizadores.

Esta experiência serviu ainda para confirmar resultados de experiências anteriores, em que se verificou que os utilizadores têm melhor desempenho, em tarefas de navegação num ambiente virtual 3D, quando utilizam o *desktop* em relação à plataforma HMD, o que pode ser eventualmente causado pela experiência com a primeira plataforma. Contudo, exprimem maioritariamente preferência pela plataforma HMD, talvez pela novidade.

Finalmente, este trabalho ilustra um caso em que se respondeu objectivamente a uma questão específica utilizando um método (experiências controladas com utilizado-

res) que, em nossa opinião, poderá ser eficaz na produção das tão necessárias directivas para o desenvolvimento de sistemas de Realidade Virtual mais usáveis.

## 6. AGRADECIMENTOS

A todos os participantes pela sua paciência e empenho e ao Dr. José Alberto Costa, professor da Escola Secundária José Estêvão de Aveiro, o convite para a participação no Dia da Ciência e Tecnologia, no âmbito do qual foi realizada a segunda etapa da experiência, bem como o apoio prestado e o entusiasmo que transmitiu aos alunos.

## 7. REFERÊNCIAS

- [Bowman04] D. Bowman, Kruijff E., LaViola Jr J. and Poupyrev I. *3D User Interfaces: Theory and practice*. Addison Wesley, (2004).
- [Conover99] W. J. Conover, *Practical non-parametric Statistics*, 3rd ed. New York: John Wiley & Sons, (1999).
- [Griffiths06] G. Griffiths, S. Sharples, J. Wilson. Performance of new participants in virtual environments: The Nottingham tool for assessment of interaction in virtual environments (NAÏVE). *Int. J. Human-Computer Studies*, 64 (2006) 240-250.
- [Hoaglin83] D., Hoaglin, F. Mosteller, J. Tukey, *Understanding Robust and Exploratory Data Analysis*, John Wiley & Sons, (1983).
- [OGRE] *OGRE: Open source 3D Graphics Engine*, <http://www.ogre3d.org/> (visitado em Maio de 2010)
- [Ruddle04] R. Ruddle, P. Péruch. Effects of proprioceptive feedback and environmental characteristics on special learning in virtual environments. *Int. J. Human-Computer Studies*, 60 (2004) 299-326.
- [Sousa-Santos09a] B. Sousa Santos, P. Dias, A. Pimentel, J. W. Baggerman, C. Ferreira, S. Silva, J. Madeira. Head Mounted Display versus desktop for 3D Navigation in Virtual Reality: A User Study. *Multimedia Tools and Applications*, 41, 1 (2009) 161-181.
- [Sousa-Santos09b] B. Sousa Santos, P. Dias, P. Santos, S. Silva, C. Ferreira, Usability evaluation in Virtual Environments through empirical studies involving users. *ACM-CHI 2009 Workshop Challenges Evaluating Usability and User Experience in Reality Based Interaction*, Boston, USA, (2009), 28-31.
- [Statistica] Statistica, [http://www.statsoftiberica.com/pt/solucoes/produtos/pacstat\\_base.html](http://www.statsoftiberica.com/pt/solucoes/produtos/pacstat_base.html) (Maio de 2010)
- [Wilson06a] J. Wilson. Interaction with Virtual environments. *Int. J. Human-Comp. Studies*, 64 (2006) 157
- [Wilson06b] Wilson, J, D'Cruz, M. Virtual and interactive environments for work of the future, *Int. J. Human-Computer Studies*, 64 (2006) 158-169.

# On the Impact of Training HRTF-Based Auralisation

Catarina Mendonça      Jorge A. Santos  
Escola de Psicologia, Univ. do Minho  
Braga  
mendonca.catarina@gmail.com  
jorge.a.santos@psi.uminho.pt

Guilherme Campos      Paulo Dias  
DETI/IEETA, Univ. de Aveiro  
Aveiro  
{guilherme.campos,paulo.dias}@ua.pt

João P. Ferreira  
Escola de Engenharia, Univ. do Minho  
Guimarães  
jpg.ferreira@gmail.com

## Abstract

*Auralisation with generic, non-individualised, Head Related Transfer Functions (HRTF) is common practice, as obtaining individualised HRTFs poses very serious practical difficulties. It is therefore extremely important to assess to what extent this hinders our 3D sound localisation capabilities. Here, we address this issue from a learning perspective. We carried out a set of experiments to better understand how the 3D virtual source localisation performance of listeners using generic HRTF is influenced by training. In experiment 1, we observed that listeners perform fairly well in terms of azimuth discrimination, but mere exposure to the tests does not cause performance improvement. In experiments 2 and 3 we implemented a short training period on both azimuth (exp2.) and elevation (exp. 3) discrimination. Training involved active learning and feedback and led to significantly better results. We therefore propose that in order to fulfil its perceptual potential, auralisation based on generic HRTF sets should always be preceded by a period of training.*

## Keywords

*Head-Related Transfer Functions, Virtual, Acoustics, Auralisation, Learning, Azimuth, Elevation.*

## 1. INTRODUCTION

Auralisation can significantly increase the feeling of immersion and improve interaction accuracy in virtual environments by providing spatial feedback and complementing visual information.

The assumption underlying earphone-based virtual acoustics is that providing a person with the same binaural stimuli he/she would get in a real environment yields the auditory perception of being in that same environment. Significant technical and scientific efforts have been carried out in recent years to create and perfect such virtual sounds. An auralised sound should provide all the necessary cues to accurately locate its virtual source. For this purpose, it is necessary to simulate room reflections, attenuation effects, interaural time and level differences (ITD and IID, respectively), as well as the shaping produced by the listener's head and pinnae. The interaural differences and other effects of the interaction of a sound wave with torso, head, pinnae (outer ears) and ear canals can be mathematically described by the binaural impulse response for the corresponding source position: the *Head Related Impulse Response (HRIR)*, or, more commonly, its Fourier transform, known as *Head Related Transfer Function (HRTF)*. Given that HRTFs depend upon ana-

tomical structures, they differ from person to person. However, given the time and effort involved in obtaining individual HRTFs, most auralisation applications rely on average sets of HRTFs measured on appropriate acoustic research manikins.

It is widely accepted that satisfactory auralisation can be obtained using average HRTFs [Loonis99]. However, there are also several reports of intracranial sound localisation and it is not yet well established how humans adapt to hearing sounds through non-individualised HRTFs.

[Wenzel93] compares localisation accuracy using external free-field acoustic sources and virtual sounds filtered by non-individualised HRTFs. This study has revealed several front-back and up-down confusions and overall similarity of the results obtained in the two test situations. On the other hand, when listeners are asked to localise complex movements, results are better with a 24-speaker system than with HRTF spatialisation [Ballas01]. Comparing the perception of sounds with individualised against generic HRTFs, the former offer significant improvement [Valjamae04].

In this paper, we address this issue taking into account the influence of learning processes. The adult human brain is able to improve its auditory localisation performance, as connectivity and response properties of the neurons are shaped by experience [King99]. There is also evidence that humans may learn to locate acoustic sources with drastically different ears [King01].

The objective of this study was to assess how training may influence the use of non-individualised HRFT. The experiments were intended to: (1) Understand the baseline localisation accuracy in subjects who had never experienced auralisation with non-individualised HRTF before; (2) Analyse the temporal evolution of localisation accuracy by simple exposure to the test sounds; and (3) Test a brief training model combining active learning and feedback.

## 2. EXPERIMENT 1

In Experiment 1, we tested the accuracy of inexperienced listeners in localising sounds at fixed elevation and variable azimuth. In 10 consecutive experimental sessions, we analysed the evolution of the subjects' performance as they became gradually more familiarised with the stimuli.

### 2.1 Method

#### 2.1.1 Subjects

Four naïve and inexperienced young adults participated in the experiment. They all had normal hearing, verified by standard audiometric screening at 500, 750, 1000, 1500 and 2000 Hz. All auditory thresholds were below 10 dB SPL and none had significant interaural sensitivity differences.

#### 2.1.2 Stimuli

The stimuli consisted of pink noise sounds. Pink noise is characterized by its power density being inversely proportional to the frequency. This provides equal energy in all octave bands, and therefore the human auditory system (which behaves analogously) perceives approximately equal magnitude in each frequency. The sounds were auralised at 8 different azimuths: 0° (front), 180° (back), 90° (left and right), (45° left and right), and 135° (left and right). They had constant elevation (0°) and distance (1m).

In the auralisation process, the original sound is convolved with the HRTF pair corresponding to the current source position. The resulting pair of signals – for the left and the right ear – is then reproduced through earphones. These HRTFs were recorded using a KEMAR dummy head microphone at the Massachusetts Institute of Technology [Gardner94]. Sounds were reproduced with a Realtec Intel 8280 IBA sound card, and presented through a set of Etymotics ER-4B MicroPro in-ear earphones.

#### 2.1.3 Procedure

All sounds were presented pseudo-randomly for 3 seconds with a 1 second interstimulus interval. There were 10 blocks of 10 stimulus repetitions each, and each block had a 5,3 minute duration. Participants were told to indicate the perceived sound source location for each stimulus.

The answers were recorded by selecting, on a touch screen, one of the eight possible stimulus positions.

## 2.2 Results and Discussion

The average percentage of correct answers for each stimulus position is presented in figure 1. As there were 8 possible answers, random answers would result in 12,5% correct answers. Thus, all results were well above chance.

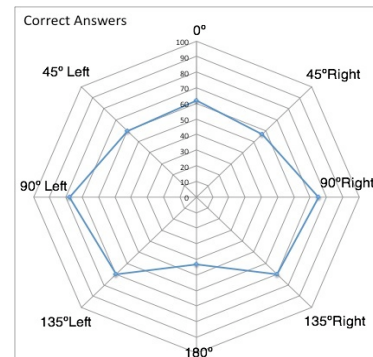


Figure 1: Percentage of correct answers for each stimulus azimuth.

As [Wenzel93] had observed, there were several front-back confusions that account for the lower accuracy at 0°, 180°, 45° and 135°. Indeed, the left and right 90° sounds were the most accurately located, with a correct response rate of 78%. The average accuracy of all azimuth localisation was significantly above chance (65%), but no ceiling performances were observed.

Analyzing the average participant's performance along time (Figure 2), we see that in spite of small fluctuations, accuracy remained largely constant.

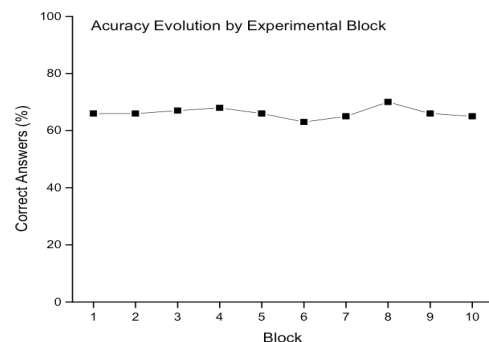


Figure 2: Average performance evolution through time

Our results reveal that naïve participants are able to discriminate sounds at several azimuths well above chance, but without ceiling performances. Throughout the exposure blocks, their accuracy does not evolve, leading to the conclusion that simple exposure is not enough for significant localisation improvement in short periods of time.

Taking these conclusions into account, a second experiment was developed where, in the same amount of time, listeners were trained to discriminate the localisation of several sounds.

## 3. EXPERIMENTS 2 AND 3

In experiment 2, we tested the participants' accuracy in localising sounds at several azimuths before and after a



short training program. In this training, we selected only a small portion of sounds and trained them through active learning and response feedback. In experiment 3, the same methodology was used, in an elevation discrimination task.

### 3.1 Method

#### 3.1.1 Subjects

Four young adults served as participants. None of them had any previous experience with virtual sounds. They all had normal hearing, tested with a standard audiometric screening, as described in experiment 1.

#### 3.1.2 Stimuli

As in experiment 1, all stimuli consisted of pink noise sounds, auralised with the same algorithms and software.

In experiment 2, the stimuli varied in azimuth, keeping elevation ( $0^\circ$ ) and distance (1m) fixed. Azimuths ranged from the front of the subjects head to their right ear, spaced at  $6^\circ$  intervals ( $6^\circ$  left,  $0^\circ$ ,  $6^\circ$  right –  $96^\circ$  right). In experiment 3, the stimuli varied in elevation, but not in azimuth ( $0^\circ$ ) or distance (1m). They ranged from the front of the listeners' head to the top in  $10^\circ$  intervals ( $0^\circ$  -  $90^\circ$ ).

All sounds had a 3 second duration, with 1 second intervals between them.

#### 3.1.3 Procedure

Both experiment 2 and 3 started with a pre-test. In the pre-test, all sounds were presented pseudo-randomly 4 times. Participants had to indicate, on a continuum displayed on a touch screen, the point in space where they expected the sound source to be.

After the pre-test, participants engaged in a training period. In experiment 2, the trained sounds corresponded to azimuths  $0^\circ$ ,  $21^\circ$ ,  $45^\circ$ ,  $66^\circ$  and  $90^\circ$ . In experiment 3, the sounds were at elevations of  $0^\circ$ ,  $50^\circ$  and  $90^\circ$ . The training conformed to the following steps:

- *Active learning:* Participants were presented with a sound player where they could hear the training sounds at their will. To select the sounds, there were several buttons on a screen, and each button clearly displayed which sound it triggered (the sound's position). The buttons were themselves arranged in the screen according to their respective position in space. For example, if the listener chose a button on the top of the screen and afterwards another button below it, he/she would hear a first sound on the top and the second sound at a lower elevation. They were informed that they had 5 minutes to learn the sounds and that afterwards they would be tested.
- *Passive Feedback:* After the 5 minutes of active learning, participants heard the training sounds and had to point their location. After each trial, they were told the correct response. Therefore, if they gave the wrong answer, they would be able to learn the correct one. The passive feedback period continued until participants could answer correctly in 80 percent of the trials (after 5 consecutive repetitions of all stimuli with at least 20 correct answers in experiment 2 and 12 correct answer in experiment 3).

After the training period, subjects performed the post-test: a testing session equal to the pre-test, to assess the discrimination differences.

### 3.2 Results and Discussion

#### 3.2.1 Experiment 2

The discrimination results of experiment 2 are presented in figure 3. The accuracy was measured as the average of the differences between the stimulus position and the response position, in azimuth degrees. The dashed line corresponds to the average error of a person responding randomly.

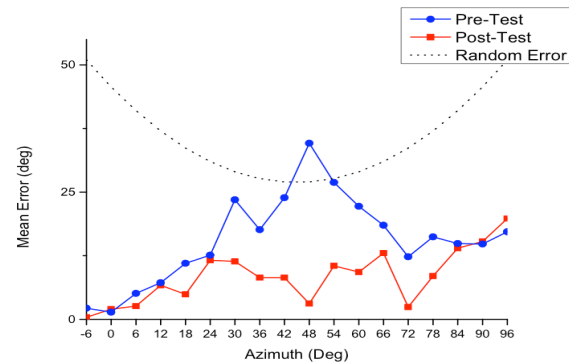


Figure 3: Average angle error in pre and post-test for azimuth

Analysing the pre-test (red) curve, we observe that azimuth discrimination is easier for frontal stimuli, where average errors are below 5 degrees. These results might be explained by the fact that there were no stimuli located at the back of the head, and therefore all front-back confusions were prevented. Similarly to the results of experiment 1, listeners were also fairly precise in identifying sound positions when these were presented laterally. On the other hand, sounds were most difficult to locate in the intermediate azimuths, between  $40^\circ$  and  $50^\circ$ . For these sounds, pre-test localisation errors were maximal. A short analysis of response accuracy along time revealed that listeners were as accurate in the beginning of the test session as in the end, confirming that simple auditory contact does not provide performance enhancement.

The training sessions were very successful for all participants. All took less than 30 minutes and in average, they lasted for 22 minutes.

The post-test results (blue curve) revealed a large error reduction ( $7.23^\circ$  in average). This difference was statistically significant in a paired samples T-test ( $t_{(287)}=14.94$ ,  $p \leq 0.001$ ). This reduction was most expressive in the intermediate azimuths, where the average error decreased 20 degrees. Analysing the trained azimuths ( $0^\circ$ ,  $21^\circ$ ,  $45^\circ$ ,  $66^\circ$ ,  $90^\circ$ ), we observe that performance enhancement was substantial not only for these stimuli, but also for others, not trained. As an example, the best error reduction was obtained with the  $48^\circ$  azimuth, a non-trained stimulus. On the other side, the  $90^\circ$  azimuth, a trained one, revealed similar results in both sessions. These findings allow us to conclude that the trained discrimination abilities for some stimuli positions are generalized to other, non-trained, auditory positions.

### 3.2.2 Experiment 3

Figure 4 displays the elevation discrimination results from the pre and post-test sessions.

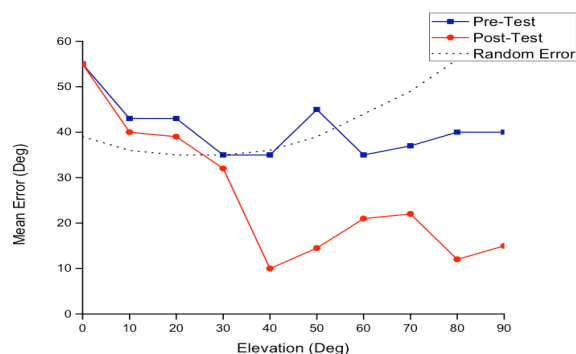


Figure 4: Average angle error in pre and post-test for elevation

Overall, participants were less accurate estimating a sound position in elevation than in azimuth. In the pre-test session, the average error was 40.8°, similar to random error. In this session, results were worst in the frontal stimuli (55° average error), but there were no large differences in acuity among all sound elevations. Again, performance was similar in the beginning and in the end of the session, confirming the absence of learning effects during exposure without feedback.

Training sessions were faster than those of experiment 1, as there were only 3 trained elevations. On average, they took 17 minutes. Only one subject did not evolve as expected. After 10 minutes testing, this subject was still making excessive mistakes, and was allowed a second learning phase, after which the 80 percent accuracy was rapidly achieved.

The post-test results were better than those of the pre-test for all subjects. This difference was significant in a paired samples T-test ( $t_{(159)}=16.678$ ,  $p<0.001$ ). The average error decreased 14.75 degrees, more than in experiment 2. The training effect was most expressive for the upper stimuli, namely at 80°, 40° and 50° elevations. Among these stimuli, the only trained one was at 50°. On the other hand, sounds at 0° elevation, a trained stimulus, revealed no decrease in the post-test session. Similarly to what was found in experiment 2, training was highly effective and well generalized to other stimuli.

## 4. FINAL DISCUSSION

In this paper we intended to analyze the accuracy of listeners in locating virtual sound sources generated with non-individualised HRTFs. We also intended to analyze the evolution of this ability along short periods of time.

In experiment 1, we assessed the acuity of the subjects in azimuth discrimination in the course of 10 testing blocks where stimuli were successively heard but no feedback was provided. The results were well above chance, but no learning effect was detected. Indeed, [King01] had reported auditory learning with severely altered ears, by mere exposure to the new sounds, but such learning took 19 days to be complete. Such a lengthy training period is impractical for most virtual sound applications.

In experiments 2 (variable azimuth) and 3 (variable elevation), we aimed at performance improvement through a training program, which combined active learning and testing with feedback. Both experiments showed there was significant performance improvement in localisation after training. This improvement was not restricted to the trained sound positions, but generalized to other source locations as well.

We conclude that in binaural auralisation using generic HRTF, it is possible to improve significantly the auditory performance of a naïve subject in a short period of time. However, this could not be achieved by mere exposure to the auralised sounds. Given the poor accuracy levels observed before training, we argue that virtual sounds with non-individualised HRTFs should only be used after short learning sessions. We propose such sessions might involve active learning, feedback and training small samples of sounds, as the new hearing abilities generalise to untrained locations.

## 5. ACKNOWLEDGEMENTS

This project was partially supported by the Portuguese Foundation for Science and Technology (SFRH/BD/36345/2007 and PTDC/TRA/67859/2006).

## 6. REFERENCES

- [Ballas01] J. A. Ballas, H. Fouad, D. Brock, J. Stroup. The effect of auditory rendering on perceived movement: Loudspeaker movement and HRTF. *Proceedings of the International Conference on Auditory Display*, 2001, Finland
- [Gardner94] B. Gardner, K. Martin. HRTF Measurements of a KEMAR Dummy-Head Microphone. url: <http://sound.media.mit.edu/resources/KEMAR.html> visited - June 2010.
- [King99] A. J. King. Auditory perception: Does practice make perfect? *Current Biology*, 9, 143-146
- [King01] A. J. King, J. W. H. Schnupp, T. P. Doubell. The space of ears to come: Dynamic coding of auditory space. *Trends in Cognitive Sciences*, 5, 261-260.
- [Loonis99] J. M. Loonis, R. L. Klatzky, R. G. Golledge. Auditory distance perception in real, virtual and mixed environments. In Y. Ohta, H. Tamura (Eds.) *Mixed Reality: Merging Real and Virtual Worlds*, Tokio: Ohmsha.
- [Valjamae04] A. Valjamae, P. Larson, D. Vastfjall, M. Kleiner. Auditory pressure, individualized Head-Related Transfer Function, and illusory ego-motion in virtual environments. *Proceedings of the Seventh Annual Workshop in Presence*, 2004, Spain
- [Loonis99] J. M. Loonis, R. L. Klatzky, R. G. Golledge. Auditory distance perception in real, virtual and mixed environments. In Y. Ohta, H. Tamura (Eds.) *Mixed Reality: Merging Real and Virtual Worlds*, Tokio: Ohmsha.
- [Wenzel93] E. M. Wenzel, M. Arruda, D. J. Kistler, F. L. Wightman. Localisation using nonindividualized Head-Related Transfer Functions. *Journal of the Acoustical Society of America*, 94, 111-123.



## **Sessão V**

### **Visualização de Informação**

Aveiro, 13, 14 e 15 de Outubro de 2010

# Exploração e Visualização de Informação Pessoal

Paulo Gomes Sandra Gama Daniel Gonçalves

Departamento de Engenharia Informática

INESC-ID / IST / Universidade Técnica de Lisboa

R. Alves Redol, 9, 1000-029 Lisboa, Portugal

paulo.gomes@ist.utl.pt, sandra.gama@ist.utl.pt, daniel.goncalves@inesc-id.pt

---

## Resumo

*Hoje em dia, a constante interacção com sistemas computacionais faz com que seja gerada uma elevada quantidade de informação pessoal. Contudo, é difícil termos uma visão global dos nossos dados e encontrar padrões na informação. Além disso, encontrar um item específico é frequentemente complexo, uma vez que os nossos dados pessoais se encontram dispersos por vários locais e aplicações. Apresentamos o VisMe, um sistema interactivo integrado para visualização de informação pessoal que permite aos utilizadores navegar e explorar os seus dados e ter uma visão global da sua vida digital. Os conceitos relevantes (pessoas, assuntos e documentos) são uniformemente mostrados em linhas temporais interligadas. Cada um destes itens pode ser progressivamente expandido em novas linhas temporais, permitindo explorar relações de forma simples e directa. Uma avaliação heurística e testes com utilizadores mostraram a viabilidade da solução apresentada como método de exploração interactiva de informação pessoal e forneceram indicações importantes para desenvolvimentos futuros.*

## Palavras-Chave

*Visualização de Informação, Gestão de Informação, Desenho Centrado no Utilizador*

---

## 1. INTRODUÇÃO

Um número crescente de dispositivos digitais faz parte das nossas vidas, permitindo-nos gerar grandes quantidades de informação pessoal, desde os documentos que escrevemos aos *e-mails* que trocamos. De forma directa ou indirecta, estes dados ajudam-nos a perceber quem somos, o que fazemos e no que estamos interessados. "O que andava a fazer em Janeiro de 2005? O João enviou-me há dois meses um artigo sobre um assunto que não me recordo. Onde encontrarei agora esse artigo?" Uma análise eficaz da nossa informação pessoal deveria esclarecer estas perguntas. Contudo, tal não é simples tendo em conta as ferramentas existentes. De facto, apesar de a capacidade de armazenamento ter vindo a crescer ao longo dos anos, os métodos e as aplicações para gerir e visualizar informação pessoal não têm sofrido melhorias substanciais.

A organização hierárquica é um dos sistemas prevalentes na organização de informação electrónica, mas sofre de algumas limitações. A principal desvantagem prende-se com o esforço permanente necessário, por parte dos utilizadores, para classificar consistentemente os dados, o que é extremamente difícil de conseguir. Além disso, diferentes tipos de informação pessoal são geridos por diferentes aplicações, com ligações escassas ou inexistentes. Como consequência, a informação pessoal encontra-se dispersa e torna-se muito difícil ter uma visão clara de como diferentes factos se interrelacionam.

Diversos sistemas têm sido desenvolvidos para visualizar diferentes tipos de colecções. No entanto, apesar de alguns destes sistemas conseguirem revelar padrões relevantes de forma atractiva e expressiva, estão limitados a fontes particulares de informação. Os que não têm essa limitação, não conseguem também fornecer uma representação unificada de informação proveniente de diversas fontes. Falta ainda, portanto, uma interface interactiva para a visualização global de toda a informação pessoal.

Propomos uma solução, a aplicação VisMe, que indexa toda a informação pessoal relevante (documentos, *e-mails*, mensagens de *instant messaging*) como um todo interrelacionado, recriando as ligações que são perdidas pelas aplicações, como entre um documento no sistema de ficheiros e o *e-mail* ao qual foi anexado. A partir desse índice, desenvolvemos uma visualização centrada nos factores autobiográficos mais relevantes: tempo, pessoas e assunto. O VisMe permite a exploração de informação pessoal de forma eficiente e compreensível. Abstraindo-se das diferentes fontes de informação, apresenta informação semanticamente relevante, e permite explorar simultaneamente, em contexto, diversas avenidas. Fornecendo uma ferramenta de visualização sinérgica, o VisMe permite aos utilizadores navegar a sua informação e ajuda-os a encontrar padrões subjectivamente relevantes.

O desenvolvimento da interface VisMe passou por ultra-

passar o desafio de representar a informação heterogénea com a qual o utilizador interage e conseguir uma forma compreensível e uniforme de visualizar essa informação. Para tal, optámos por um desenho iterativo da interface, desde a criação de protótipos de baixa fidelidade, até ao protótipo final, descrito na Secção 3. O protótipo foi validado através de avaliação heurística, o que permitiu refiná-lo antes de partir para a validação com utilizadores, garantindo assim que a interface permite, de facto, explorar e visualizar a informação pessoal do utilizador de forma interactiva e significativa para o utilizador.

## 2. TRABALHO RELACIONADO

Uma vez que a visualização de informação pessoal consiste num método com imenso potencial para gestão de informação pessoal, têm sido desenvolvidas diversas aplicações neste âmbito, normalmente focando-se numa única fonte de informação, como *e-mails*, registos de aplicações de *instant messaging* ou documentos de texto.

No campo da visualização de *e-mails*, o *Themail* [Viegas 06] distingue-se pela interface simples e apelativa e pela capacidade de mostrar padrões no conteúdo dos *e-mails*. Utilizando simples linhas temporais, revela o conteúdo de colecções de *e-mails* e a sua evolução ao longo do tempo. O *ChristalChat* [Tat 06], uma visualização para *instant messaging*, revela as trocas de informação com diversos contactos num espaço conversacional pessoal através de estrutura tridimensional, mas não representa o conteúdo para além de um indicador periférico de humor através da cor. Já *ThemeRiver* [Havre 00] consegue representar visualmente a evolução de diferentes tópicos contidos numa colecção de documentos de texto sob a forma de correntes coloridas que crescem e diminuem ao longo do tempo.

Também têm sido desenvolvidos sistemas com o intuito de visualizar informação de diversas fontes. O *Milestones in Time* [Ringel 03] utiliza uma visualização familiar em listas, a que junta uma linha temporal preenchida com pontos de referência, fornecendo uma interface atractiva para procura de um historial de informação multimédia. O *FacetMap* [Smith 06] consegue juntar uma representação visual das facetas (sob a forma de bolhas) com o próprio mecanismo de procura de uma forma simples. O *Feldspar* [Chau 08], por sua vez, permite aos utilizadores construir perguntas de associação de forma interactiva e incremental. Os seus pontos fortes estão relacionados com a associação entre entidades que o utilizador pode usar para encontrar informação. Foca-se mais nas ligações entre entidades do que nas entidades em si, tornando possível encontrar mais dados sobre as entidades do que procurando separadamente por cada um desses itens.

No entanto, nenhum destes métodos de visualização fornece uma visão global unificada do conteúdo de uma colecção heterogénea de documentos. O VisMe tenta colmatar esta lacuna através de uma visualização interactiva de informação pessoal proveniente de diversas fontes, permitindo encontrar padrões relevantes para o utilizador.

## 3. SOLUÇÃO PROPOSTA

Para explorar a colecção pessoal de documentos de um utilizador, o primeiro passo é encontrar uma forma de recolher e indexar essa informação. Para tal, foi utilizado o Scribe [Gonçalves 08], uma ferramenta que indexa e inter-relaciona *e-mails*, documentos, *logs* de *instant messaging*, páginas web, etc. Sobre estes dados indexados, foi desenvolvida uma camada para facilitar a integração e fornecer acesso eficiente à informação pessoal. Depois de ter acesso aos dados pessoais dos utilizadores, desenvolvemos uma interface que lida com a representação da informação e com a interacção com o utilizador.

Focámo-nos em três objectivos ao conceber esta interface: (1) Que fosse de simples compreensão e manipulação, (2) Que permitisse explorar os dados em contexto, como um todo interrelacionado; e (3) Que lidasse uniformemente com dados de várias fontes. Através de esboços e prototipagem iterativos, chegámos à solução proposta. A ideia fundamental é que cada elemento na visualização, nomeadamente palavras-chave (palavras mais significativas de cada documento), contactos (autores, remetentes e destinatários de informação) e documentos (ficheiros, *e-mails* e *logs* de *instant messaging*) podem ser expandidos para mostrar as palavras-chave, os contactos e os documentos com que, por sua vez, se relacionam (todos os documentos de um autor, todas as palavras-chave num documento, todas as palavras-chave em mensagens de e para uma pessoa em particular, etc.). Cada elemento na visualização é representado por uma palavra e três botões a partir dos quais várias linhas podem surgir, uma por faceta (Figura 1).

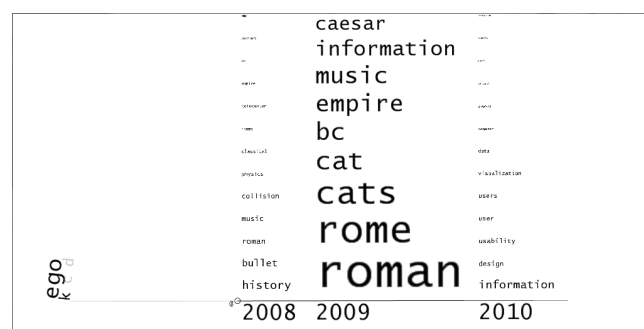


Figura 1. Palavras-chave expandidas.

A visualização começa com o elemento "ego", que representa o utilizador. Ao colocar o cursor sobre este elemento, surgem três botões que representam as palavras-chave, os contactos e os documentos ("k", "c", e "d"). Ao clicar num destes, é criada uma linha temporal que representa os elementos respectivos. Os elementos mais representativos aparecem maiores e mais perto do fundo. Além de ser possível expandir uma linha temporal verticalmente com um único clique, pode arrastar-se a mesma a partir do ícone para qualquer posição e orientação.

As linhas temporais mostram inicialmente os elementos mais relevantes de todos os tempos numa única linha ho-

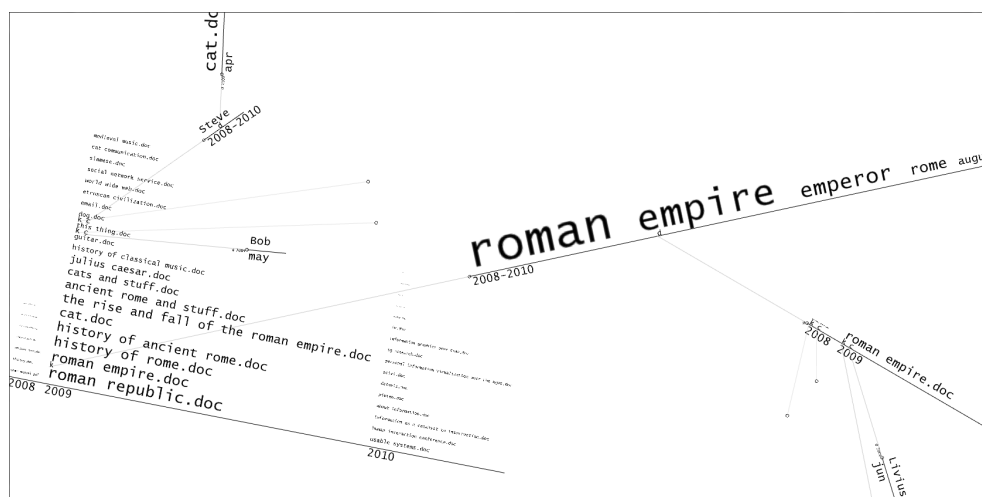


Figura 2. Linhas temporais expandidas a partir de vários elementos.

rizoidal, sendo depois possível ver progressivamente anos, meses, e dias clicando no período desejado. Por exemplo, seleccionar um ano modificará a linha temporal para mostrar os meses respectivos.

No seguimento de uma expansão, qualquer elemento da linha temporal, seja uma palavra-chave, um contacto ou um documento, pode ser submetido ao mesmo processo que o elemento inicial e assim sucessivamente. O utilizador pode explorar progressivamente a informação apresentada através das linhas temporais.

O VisMe pode manter vários elementos expandidos no mesmo ecrã (Figura 2), que pode sofrer translação, rotação e escala manualmente por parte dos utilizadores. Estes podem ainda fazer duplo clique numa linha temporal para aí focar a visualização através de rotação, escala e translação automáticas.

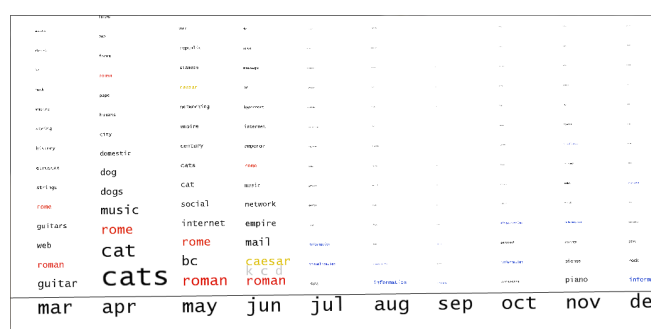


Figura 3. Elementos destacados com cores.

Quando os utilizadores movem o rato sobre um elemento, todas as instâncias desse elemento, em qualquer linha temporal, ficam destacadas (Figura 3), facilitando o acompanhamento da evolução ao longo do tempo. A cor depende da posição do rato sobre o elemento, seguindo um gradiente do vermelho ao azul, passando pelo verde. Ao clicar, a cor é fixada, sendo que um novo clique restabelece a cor preta.

Optámos por este método pela sua simplicidade, uma vez que permite a selecção de várias cores distintas sem sobrecarregar a interface com botões ou outros elementos adicionais. Se for necessário aplicar precisamente a mesma cor a vários elementos (para destacar diferentes grupos de contactos, por exemplo), é possível copiar a cor de qualquer elemento usando sobre este o botão direito do rato. Depois, ao passar o rato sobre outro elemento, este terá a cor seleccionada. Para soltar esta cor, basta um clique com o botão direito num espaço vazio.



Figura 4. Palavras-chave parcialmente destacadas pela procura por texto.

O VisMe suporta também pesquisa por texto. Ao premir uma letra no teclado, uma caixa de *input* é mostrada no topo do ecrã, contendo o texto à medida que este é escrito. Sempre que uma letra é inserida, o texto resultante é procurado e os resultados são mostrados ao utilizador através de possíveis correspondências a cinzento e de uma indicação do número de resultados para a faceta actualmente seleccionada. Diferentes possibilidades de correspondência podem ser seleccionadas carregando nas teclas de cima e baixo, sendo possível completar o texto introduzido de acordo com a possibilidade escolhida carregando na tecla de *tab*. Se o texto estiver presente

nas linhas temporais expandidas, os elementos correspondentes serão destacados (Figura 4). Se um dado elemento se incluir na procura mas não for suficientemente representativo para ser mostrado na linha temporal, aparecerá no topo da coluna respectiva, significativamente maior do que o elemento anterior (Figura 5), o que mostra que este não segue a mesma convenção de tamanho que o resto da linha temporal. Além disso, haja ou não uma correspondência exacta em qualquer período temporal, o tempo correspondente será destacado.



**Figura 5. Correspondência exacta com o texto procurado faz emergir um documento.**

Evidentemente, existe um limite para a quantidade de informação que pode ser mostrada de forma compreensível no mesmo ecrã. Assim sendo, implementámos diversas medidas que visam maximizar a quantidade de informação disposta simultaneamente e minimizar o ruído visual. Quando o rato não está sobre uma linha temporal, apenas os elementos mais representativos, localizados na base ao longo da linha, são mostrados. Além disso, as linhas temporais inicialmente aparecem num estado horizontal, com os elementos mais representativos para todos os períodos temporais. Assim, grande parte do tempo, são apenas visíveis linhas relativamente estreitas, correspondendo aos elementos mais importantes. Permitimos também aos utilizadores reposicionar ou esconder linhas temporais individualmente, clicando ou arrastando um botão circular na sua base. Movendo a vista e reorganizando a informação, os dados relevantes são mantidos em vista, sendo sempre possível recuperar todo o contexto através de *zoom out* e *panning*.

#### 4. AVALIAÇÃO HEURÍSTICA

Um primeiro protótipo da solução descrita acima foi submetido a uma avaliação heurística com o propósito de identificar problemas de usabilidade na interface com o utilizador.

##### 4.1. Método

Foi pedido a quatro peritos que executassem três tarefas construídas com o intuito de exercitar todas as funcionalidades do protótipo e que avaliassem a interface segundo o

cumprimento das heurísticas de Nielsen.

Os dados utilizados foram os mesmos para os quatro analistas e consistiram numa colecção de 106 documentos de texto da autoria de 13 pessoas distintas ao longo de um espaço de 3 anos.

A primeira tarefa envolvia a procura de um documento com base no conhecimento do seu autor, tema principal, e data, tendo o propósito de avaliar as funcionalidades básicas de procura e recuperação de documentos. A segunda misturava o reconhecimento de um padrão, a maior relevância de um determinado tema num período de tempo a determinar, com a procura de um documento sobre outro tema nesse período, de modo a testar a viabilidade de um dos modos de utilização considerados na origem deste projecto (procura de um documento com base no contexto temporal ligado a um tema). A terceira requeria que os analistas descrevessem a evolução de um tema e de que forma seria influenciada pela evolução de outro tema num certo período de tempo, para, desta forma, verificar se a visualização permite discernir quais os elementos mais importantes e como estes evoluem temporalmente. Os peritos tiveram uma hora para trabalhar nestas tarefas e tomar nota, por escrito, de todos os problemas com a interface, incluindo observações e sugestões de resolução. Depois da leitura desses comentários, houve ainda um período de cerca de trinta minutos em que conversamos com os analistas de modo a clarificar algumas dessas observações e sugestões.

##### 4.2. Resultados

Heurística	Cumpre	Não Cumpre
Estado do sistema visível	2	2
Linguagem do utilizador	3	1
Controlo; livre arbítrio	3	1
Consistência e normas	4	0
Evitar erros	2	2
Reconhecimento	0	4
Flexibilidade e eficiência	4	0
Desenho estético e minimalista	3	1
Reconhecer e recuperar de erros	1	3
Ajuda e documentação	0	4

**Tabela 1. Cumprimento das heurísticas de Nielsen de acordo com cada um dos quatro peritos.**

A avaliação heurística identificou vários problemas com a interface. Na tabela 1 é possível identificar o cumprimento das várias heurísticas consideradas pelos quatro peritos.

A falta de visibilidade do estado do sistema foi associada à navegação, especificamente à tendência de alguns dos peritos de se perderem após a expansão e manipulação de linhas temporais durante longos períodos de tempo. Um perito sugeriu a implementação de uma bússola que apontasse para o centro ou para o início da visualização. Uma

solução mais directa pode consistir em mover e escalar a vista do sistema de forma a visualizar toda a estrutura da visualização no ecrã, isto de forma gradual e suave após a activação de um botão ou o premir de uma tecla.

A falta de prevenção de erros observada por dois dos peritos deveu-se ao facto de ser permitido abrir várias linhas temporais umas em cima das outras, causando a sobreposição de botões e outros elementos, e também à falta de distinção clara entre elementos interactivos e não interactivos. O segundo problema foi resolvido alterando o texto dos elementos para negrito quando o rato se encontra em cima destes e é possível clicar para realizar qualquer acção. Uma possível solução para o primeiro problema está a ser trabalhada actualmente e baseia-se na detecção de colisões com separação suave de linhas temporais.

A falta de funcionalidades para facilitar o reconhecimento em vez de lembrança foi apontada por todos os peritos. A omissão de informação referente ao significado dos botões e dos diferentes tamanhos dos elementos nas linhas temporais, por exemplo, podia ser tratada com *tooltips* e com indicações visuais claras, apesar de não ter havido nenhuma sugestão sobre esta segunda possibilidade.

Em caso de erro, como um bug em concreto que levava ao desaparecimento de uma linha temporal quando os analistas tentavam progredi-la para mostrar os dias de um determinado mês, o protótipo não fornecia qualquer indicação do sucedido. Foi sugerido que uma simples caixa de texto com descrições claras sobre o problema e indicações simples sobre a sua resolução deveria ser aplicada nestes casos. Um perito chamou também a atenção para a falta de *undo* e *redo*.

A lacuna em termos de documentação e ajuda foi apontada claramente por todos os peritos. *Tooltips*, já sugeridas para auxiliar ao reconhecimento, podiam também contribuir neste caso. Um perito sugeriu também um ecrã de ajuda em que seriam explicados sucintamente todos os elementos da visualização e todas as funcionalidades disponíveis.

Apesar destes problemas, os peritos consideraram o sistema esteticamente agradável, de compreensão simples, e fácil de usar, precisamente algumas das propriedades que pretendíamos atingir ao desenvolver esta interface.

## 5. TESTES COM UTILIZADORES

Entre a avaliação heurística e os testes com utilizadores, resolvemos uma série de erros que se apresentavam no protótipo. Implementámos também um mecanismo de procura por texto, descrito anteriormente. Para validar a interface, considerámos necessário utilizar uma colecção de documentos realista, e isso implicava uma grande dificuldade em encontrar palavras-chave, contactos, e nomes de documentos sem esta funcionalidade.

### 5.1. Método

Pedimos a 20 voluntários com idades entre os 17 e os 29 anos ( $\bar{x} = 23.7$ ,  $\sigma = 2.7$ ,  $\tilde{x} = 23.5$ ), auto-avaliados com

um elevado nível de experiência informática ( $\bar{x} = 3.7$ ,  $\sigma = 0.47$ ,  $\tilde{x} = 4$ ,  $IIQ = 1$ , escala de 1 a 4) que desempenhassem uma série de tarefas, usando o VisMe, num conjunto de 1004 documentos de texto produzidos por 102 pessoas no espaço de 3 anos. Este conjunto de dados foi produzido de acordo com os nossos conhecimentos acerca de espaços de informação pessoal, para que fosse representativo de uma colecção real de documentos, com tendências e padrões realistas para cada combinação testada de facetas e documentos, além de autores suficientes para que fosse difícil encontrá-los acidentalmente sem ajuda da interface.

Foram testadas oito tarefas de procura de documentos, que consistiram na recuperação de um documento com o conhecimento de uma ou mais facetas (o tempo, a palavra-chave mais representativa e o autor, bem como uma tarefa única em que o nome do ficheiro era dado) e nove tarefas de detecção de padrões e tendências, que envolviam a identificação de elementos mais representativos de um determinado período temporal (um ano) e vice-versa, bem como a descrição de como esses elementos evoluíam ao longo desse período (em que mês aparecia, em que mês desaparecia, se a sua representatividade tinha crescido ou decrescido entre tanto). Estas tarefas foram executadas por uma ordem semi-aleatória: dez utilizadores realizaram primeiro as tarefas de recuperação de documentos, os outros dez começaram pelas tarefas de detecção de padrões. Dentro de cada grupo de tarefas, estas foram ordenadas aleatoriamente.

Antes de serem entregues as tarefas aos utilizadores, foi feita uma demonstração de cinco minutos do protótipo, após a qual os utilizadores dispuseram de outros cinco minutos para explorar livremente a interface. Nestes dois casos, foram utilizados dados distintos dos usados nas tarefas para não influenciar a sua execução.

As tarefas foram cronometradas e gravadas para análise. Foi pedido aos utilizadores que, depois de executarem cada tarefa, classificassem o grau de dificuldade da mesma, numa escala de 1 (muito difícil) a 4 (muito fácil). O tempo limite para cada tarefa era de 150 segundos, após o qual era pedido aos utilizadores que seguissem para a tarefa seguinte.

No final de cada sessão, foi pedido aos utilizadores que preenchessem um questionário para avaliar a sua satisfação relativamente à interface e a algumas funcionalidades específicas, através de uma série de itens a classificar numa escala de quatro pontos (por exemplo, "Aprendizagem do sistema: difícil 1 2 3 4 fácil").

### 5.2. Resultados

Depois de realizados os testes com os utilizadores, procedemos a uma análise estatística dos dados obtidos. Apresentamos de seguida a média aritmética ( $\bar{x}$ ), o desvio padrão ( $\sigma$ ), e a mediana ( $\tilde{x}$ ) de diversas variáveis. Foi também utilizado o intervalo inter-quartil (IIQ) no caso dos dados extraídos dos questionários.



### 5.2.1. Procura de documentos

A maior parte das tarefas de procura de documentos foi completada com sucesso (com falhas por tarefa nas vinte sessões  $\bar{x} = 2$ ,  $\sigma = 2.4$ ,  $\tilde{x} = 1$ ). Em média, as tarefas bem sucedidas foram completadas em cerca de 52 segundos ( $\bar{x} = 52.2$ ,  $\sigma = 27.1$ ,  $\tilde{x} = 45.5$ ). Os utilizadores consideraram também que as tarefas eram fáceis de completar ( $\bar{x} = 3.4$ ,  $\sigma = 0.899$ ,  $\tilde{x} = 4$ ), o que mostra que o VisMe é uma ferramenta eficaz para recuperar documentos por tempo, palavras-chave e autores.

Existem, contudo, duas excepções. Ambas as tarefas em que era pedido aos utilizadores que localizassem um documento baseando-se numa combinação de palavra-chave e autor (e, numa das tarefas, também tempo) tiveram um maior número de falhas (6 e 5 em 20 sessões), os maiores tempos de realização ( $\bar{x} = 84.5$ ,  $\sigma = 27.5$ ,  $\tilde{x} = 77$  e  $\bar{x} = 88.7$ ,  $\sigma = 27.6$ ,  $\tilde{x} = 84$ ) e foram consideradas as tarefas mais difíceis ( $\bar{x} = 2.6$ ,  $\sigma = 1.14$ ,  $\tilde{x} = 3$  e  $\bar{x} = 2.6$ ,  $\sigma = 0.94$ ,  $\tilde{x} = 2.5$ ). É possível completar estas tarefas, como a maior parte dos utilizadores fez, mas é evidentemente mais difícil encontrar documentos com base na combinação de duas facetas (excepto o tempo).

Realizar estas tarefas requer expandir documentos de ambas as facetas e comparar os resultados, ou expandir documentos de uma faceta e expandir a outra faceta a partir de cada um desses documentos, um por um. Com o conjunto de dados utilizado, o número de documentos correspondente a uma das facetas (ou uma das facetas mais o tempo) vai desde dois a doze, tornando exequível a verificação tanto de cada item individual como dos resultados em duas linhas temporais separadas. Conjuntos de resultados mais extensos podiam ter tornado estas tarefas impraticáveis. Alguns utilizadores exprimiram alguma frustração relativamente ao facto de as linhas temporais não serem filtradas de acordo com a sua hierarquia, já que expandir documentos de uma palavra expandida de um contacto não resulta numa lista de documentos escritos apenas por esse autor acerca dessa palavra-chave, o que, a acontecer, simplificaria esta tarefa.

Outros valores que se destacam são os da combinação de uma palavra-chave ou de um contacto com uma data. Estas tarefas demoraram ligeiramente mais ( $\bar{x} = 48.4$ ,  $\sigma = 18.6$ ,  $\tilde{x} = 44.5$  e  $\bar{x} = 58.6$ ,  $\sigma = 19.03$ ,  $\tilde{x} = 51$ ) do que localizar um documento a partir de uma palavra-chave ( $\bar{x} = 36.15$ ,  $\sigma = 14.56$ ,  $\tilde{x} = 35.5$ ) e de um contacto ( $\bar{x} = 34.6$ ,  $\sigma = 12.8$ ,  $\tilde{x} = 34.5$ ), e não foram bem sucedidas duas e três vezes, respectivamente, apesar de não serem consideradas muito mais difíceis ( $\bar{x} = 3.5$ ,  $\sigma = 0.82$ ,  $\tilde{x} = 4$  e  $\bar{x} = 3.25$ ,  $\sigma = 0.91$ ,  $\tilde{x} = 3.5$ ) do que a média.

Alguns utilizadores tentaram primeiro localizar os documentos apenas através da data, voltando depois para trás para encontrar os documentos relacionados com a outra faceta dada. Mesmo não cometendo este erro, é compreensível que a acção de aproximação temporal carregando sucessivamente nas datas de uma linha temporal adicionem alguns segundos ao tempo de execução destas tarefas. O facto de demorar mais tempo a localizar um do-

cumento com base apenas numa data ( $\bar{x} = 47.6$ ,  $\sigma = 19.4$ ,  $\tilde{x} = 39$ ) do que com base apenas numa palavra-chave ( $\bar{x} = 36.15$ ,  $\sigma = 14.6$ ,  $\tilde{x} = 35.5$ ), contacto ( $\bar{x} = 34.6$ ,  $\sigma = 12.9$ ,  $\tilde{x} = 34.5$ ), e nome de ficheiro ( $\bar{x} = 38.6$ ,  $\sigma = 20.3$ ,  $\tilde{x} = 32.5$ ) suporta esta observação.

Por fim, apesar de os tempos de recuperação serem, aparentemente, mais elevados do que os tempos possíveis com tradicionais ferramentas de pesquisa baseadas em palavras-chave, a estas últimas falta o suporte para algumas tarefas, mais complexas, como as que são requisitadas aos utilizadores do VisMe. Por exemplo, encontrar um documento com base na combinação do autor, da data e do assunto, obrigaria os utilizadores a pesquisar por várias palavras-chave e inter-relacionar a informação por si mesmos, aumentando o tempo necessário para a tarefa. O VisMe fornece suporte explícito para essas tarefas. Apesar, de, como descrito, estas serem as tarefas em que o VisMe não tem um desempenho tão satisfatório, a realização das tarefas foi elevada, mostrando que é útil, já que foi desenhado precisamente para mostrar múltiplas avenidas de exploração de cada vez, ajudando o utilizador a recordar-se do contexto dos documentos. Melhorias futuras na interface (ver abaixo) aumentarão o desempenho nestas tarefas mais complexas.

### 5.2.2. Detecção de padrões e tendências

As tarefas de detecção de padrões foram concluídas com sucesso pela maior parte dos utilizadores (com falhas por tarefa nas vinte sessões  $\bar{x} = 0.44$ ,  $\sigma = 0.53$ ,  $\tilde{x} = 0$ ) e foram consideradas muito fáceis ( $\bar{x} = 3.78$ ,  $\sigma = 0.52$ ,  $\tilde{x} = 4$ ). Em média, as tarefas bem sucedidas foram concluídas em cerca de 35 segundos ( $\bar{x} = 35.49$ ,  $\sigma = 18.99$ ,  $\tilde{x} = 30$ ). Não existe uma tarefa específica que se destaque, pelo que se conclui que a representação unificada de elementos em linhas temporais permite detectar palavras-chave e contactos mais importantes e apreciar correctamente a sua evolução ao longo do tempo com uma facilidade uniforme.

É, no entanto, possível salientar um pormenor observado em algumas tarefas que envolviam a descrição da evolução de uma palavra-chave ou de um contacto num determinado ano. Notou-se, nestes casos, uma fraca utilização do destaque de elementos de uma linha temporal através da cor, uma funcionalidade criada especificamente para facilitar a visão da evolução de elementos ao longo do tempo. Muitos dos utilizadores limitavam-se a observar as palavras sem qualquer destaque, outros faziam-no apenas momentaneamente passando o rato por cima. Uma situação interessante foi a utilização da procura por texto por alguns utilizadores. Ao escrever um contacto ou uma palavra-chave, este elemento fica demarcado de forma mais pronunciada do que o destacamento simples com a passagem do rato, marcando até os períodos temporais em que os elementos ocorrem e facilitando bastante estas tarefas. A partir do momento em que se aperceberam disso, foi comum os utilizadores recorrerem a essa funcionalidade nas tarefas seguintes. O impacto do destaque com cores,

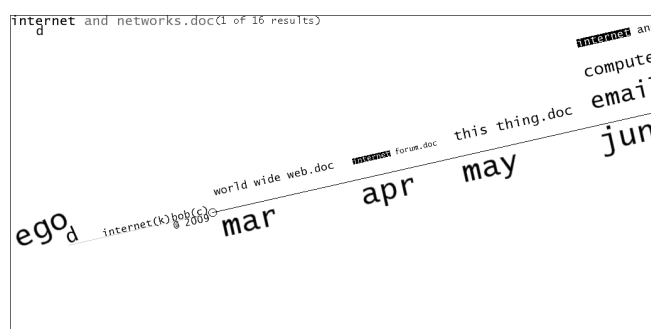
pelo contrário, não foi tão forte ao ponto de levar os utilizadores a valerem-se dele de forma pronunciada. No entanto, o destaque com procura por texto só pode ser aplicado a um elemento de cada vez, limitação essa que não foi evidente nestes testes dado não ser necessário demarcar vários elementos simultaneamente em cada tarefa. Pode, no entanto, ser reconsiderada a forma como é feito o destaque com a passagem do rato e torná-lo tão forte como o da procura por texto.

### 5.2.3. Questionário

O questionário mostrou alguma satisfação com o sistema por parte dos utilizadores ( $\bar{x} = 3.35$ ,  $\sigma = 0.49$ ,  $\tilde{x} = 3$ , IIQ = 1). Estes não o acharam difícil de usar ( $\bar{x} = 3.20$ ,  $\sigma = 0.69$ ,  $\tilde{x} = 3$ , IIQ = 1), mas por vezes acharam-no algo difícil de aprender ( $\bar{x} = 2.75$ ,  $\sigma = 0.85$ ,  $\tilde{x} = 3$ , IIQ = 1). Sentiram também que nem sempre o sistema oferecia funcionalidades suficientes ( $\bar{x} = 2.85$ ,  $\sigma = 0.59$ ,  $\tilde{x} = 3$ , IIQ = 0.25). Apesar de os utilizadores não considerarem, em geral, difícil o controlo da área visível através de rotação, escala e translação ( $\bar{x} = 3.15$ ,  $\sigma = 0.81$ ,  $\tilde{x} = 3$ , IIQ = 1.25), observámos que muitos utilizadores se mostravam desconfortáveis com estas acções. Este facto pode ser explicado pela curta experiência com uma interface desconhecida e não convencional.

## 6. TRABALHO FUTURO

Os testes realizados vieram a mostrar que existe um problema com a recuperação de um documento baseado na combinação de duas facetas. Pode esperar-se que este problema se agrave se os utilizadores tentarem uma combinação de diversos autores e palavras-chave. Uma solução possível para este problema pode ser a filtragem.



**Figura 6. Todos os documentos escritos pelo Bob contendo palavra Internet.**

Desenvolvemos uma solução funcional, embora ainda não testada. No protótipo actual, os utilizadores podem simplesmente arrastar uma palavra-chave, contacto e nome de ficheiro, para qualquer linha temporal, em qualquer combinação e as vezes que desejarem, para executarem a filtragem. Clicando e arrastando o rato de uma palavra-chave, por exemplo, numa linha temporal para o espaço

ocupado por uma segunda linha temporal acrescentará a palavra-chave como filtro desta última. Os filtros activos aparecem à esquerda da linha temporal e podem ser removidos através de um simples clique. Este conceito foi implementado também no texto de pesquisa que aparece no canto do ecrã, que também pode ser arrastado para qualquer linha temporal, tornando extremamente fácil a filtragem de uma linha temporal de acordo com qualquer faceta que os utilizadores encontrem através de procura textual.

Planeamos também modificar o protótipo para que os filtros passem através das hierarquias de linhas temporais, correspondendo à expectativa demonstrada por alguns utilizadores. Uma vez que é ainda importante aceder a toda a informação relativa a qualquer elemento, em qualquer lado, arrastar linhas temporais com o botão direito do rato expandirá linhas temporais não filtradas.

Apesar de não se ter tornado especificamente evidente nos testes com utilizadores (a quantidade de informação necessária em ecrã para cada tarefa era relativamente pequena), estamos a investigar soluções adicionais para sobrecarregamento de elementos visuais, como a já referida detecção de colisões entre linhas temporais.

Considerando os resultados da avaliação heurística e o desconforto observado em alguns utilizadores durante os testes, estamos também a estudar formas de melhorar o controlo do campo de visão da visualização. Isto poderá passar por um controlo automático e opcional da mesma de modo a mostrar sempre a totalidade dos dados explorados e pela suavização dos controlos manuais de modo a tornar a sua utilização menos intimidante.

## 7. CONCLUSÃO

Um método eficiente e integrado de visualização de toda a informação pessoal contida nos nossos computadores pode ser a solução para procurar e recuperar ficheiros, descobrir padrões interessantes, ou simplesmente desfrutar da exploração da nossa vida digital. Apresentámos a nossa solução, o sistema interactivo de visualização pessoal VisMe. A principal ideia subjacente ao VisMe é a de expandir progressivamente e dispor temporalmente a informação relativa a uma de três facetas: palavras-chave, pessoas e assunto. Esta solução apresenta a vantagem de fornecer uma representação unificada e coerente de informação heterogénea, além de mostrar uma visão global do conteúdo de uma colecção de dados de uma forma que potencia o inter-relacionamento dos mesmos.

Uma avaliação heurística e testes com utilizadores validaram as funcionalidades básicas desta interface para a procura de documentos e detecção de padrões, mas também mostraram alguma fragilidade na combinação de várias facetas na mesma pesquisa, algo que tentamos solucionar posteriormente com desenvolvimento de um mecanismo de filtragem que consegue tirar partido da presença de múltiplas linhas temporais e da pesquisa por texto.

No entanto, porque os testes não foram realizados com os

dados dos próprios utilizadores, não validámos completamente a nossa solução. Idealmente, o VisMe utiliza a memória que um indivíduo tem acerca da sua informação pessoal; o contexto em torno de cada documento, a história de cada contacto, etc. Não houve esse auxílio nestes testes. Os dados artificiais também inviabilizaram a comparação do desempenho do VisMe com abordagens tradicionais, como navegar pastas de documentos ou utilizar mecanismos de pesquisa. Ainda assim, o objectivo foi validar os progressos iniciais no desenvolvimento da nossa solução e estamos já a planear realizar testes com dados reais de cada utilizador em que o VisMe será também comparado com outras abordagens de procura de documentos.

## Referências

- [Chau 08] D. H. Chau, B. Myers, e A. Faulring. Feldspar: A system for finding information by association. Em *ACM SIGCHI PIM2008, the Third International Workshop on Personal Information Management, Florence, Italy*, 2008.
- [Gonçalves 08] D. Gonçalves e J. Jorge. In search of personal information: narrative-based interfaces. Em *IUI 2008, New York, NY, USA*, páginas 179–188, 2008.
- [Havre 00] S. Havre B. Hetzler e L. Nowell. The-meriver: Visualizing theme changes over time. Em *Proc. IEEE Symposium on Information Visualization*, páginas 115–123, 2000.
- [Ringel 03] M. Ringel, E. Cutrell, S. T. Dumais, e E. Horvitz. Milestones in time: The value of landmarks in retrieving information from personal stores. Em *INTERACT*, 2003.
- [Smith 06] G. Smith, M. Czerwinski, B. Meyers, D. Robbins, G. Robertson, e D.S. Tan. Facetmap: A scalable search and browse visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):797–804, 2006.
- [Tat 06] A. Tat e S. Carpendale. Crystalchat: Visualizing personal chat history. *Hawaii International Conference on System Sciences*, 3:58c, 2006.
- [Viegas 06] F. Viegas, S. Golder, e J. Donath. Visualizing email content: portraying relationships from conversational histories. Em *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, páginas 979–988, New York, NY, USA, 2006. ACM.

# Metabrain – Conhecimento na era do PetaByte

João Teixeira

Gabriel Barata

Daniel Gonçalves

Dep. Eng<sup>a</sup>. Informática, IST  
Av. Rovisco Pais, 1000 Lisboa

{joao.teixeira,gabriel.barata}@ist.utl.pt, daniel.goncalves@inesc-id.pt

## Sumário

*Hoje em dia, a Internet é uma fonte enorme de informação sobre os mais diversos ramos do conhecimento. No entanto, este conhecimento encontra-se disperso por muitos sítios, sem qualquer ligação ente eles, o que torna difícil inter-relaciona-lo e compreendê-lo. O objectivo deste trabalho é identificar e avaliar métodos de extracção de informação de forma simples e eficaz, sem recorrer a parsing de língua natural ou bases complexas de conhecimento anotado. Queremos mostrar que é possível extrair informação de forma implícita na Web através de métodos estatísticos. Para além disso, dados em bruto são normalmente de difícil compreensão. Como tal, procurámos também estudar como permitir aos utilizadores usar técnicas de visualização de informação de forma simples e eficaz para melhor os analisar e entender. Neste trabalho, propomos uma nova abordagem onde os utilizadores podem criar os seus próprios extractores de informação e respectivas visualizações, sem precisar de escrever uma única linha de código, de uma maneira fácil e altamente flexível, através de uma interface especialmente desenvolvida para o efeito. Um protótipo deste sistema, MetaBrain, foi desenvolvido e encontra-se em fase de testes e afinações.*

## Palavras-chave

*Extracção de conhecimento, inferência probabilística, data mining, visualização, Internet.*

## 1. INTRODUÇÃO

A versatilidade da Internet é também o seu Calcanhar de Aquiles. Qualquer pessoa pode criar páginas das mais díspares formas e conteúdos. Isto permitiu o seu rápido crescimento e que hoje em dia seja um repositório enorme de informação. No entanto, a Internet é um meio muito pouco estruturado, tornando difícil a extracção de informação. Apenas a leitura por um humano permite efectuar esta tarefa com algum grau de sucesso e, mesmo assim, a dispersão da informação obriga a consultar muitas fontes, nem sempre consistentes.

Para além disto, há todo um conjunto de informação não explícita ao nosso dispor. As páginas e os seus conteúdos reflectem um "inconsciente colectivo". Os temas versados e a forma em como o são revelam tendências e padrões que podem não ser imediatamente aparentes.

Tenta-se recorrer a ferramentas relativamente sofisticadas para fazer face a toda esta complexidade e extrair informação semanticamente relevante da Web. Desde processamento de língua natural a redes semânticas, de tudo se tem usado. Mas essas técnicas são frágeis e os resultados obtidos são válidos apenas em contextos muito restritos. A necessidade de uma base de conhecimento bem estruturada e anotada para o bom funcionamento deste tipo de análises é outro dos grandes problemas.

Existe no entanto, outra forma de abordarmos o problema. A utilização de métodos de análise estatística permitem-nos extrair informação relevante da Web. Por exemplo, o simples número de resultados de uma pesquisa no Google fornece informação, sem que seja necessário

compreender o conteúdo das páginas retornadas. Imagine-se que queríamos saber se o Barak Obama é ou não a favor do encerramento da prisão de Guantanamo Bay. "Barack Obama Guantanamo keep open" devolve 96500 resultados e "Barack Obama Guantanamo close" 1230000. É bastante evidente qual a sua opinião sobre esse assunto, e não foi preciso fazer *parsing* de nada nem compreender língua natural.

Assim, o objectivo deste trabalho é estudar a melhor forma de explicitar o conhecimento que existe na Internet de forma implícita, fazendo-o de forma robusta e independente do domínio. Por outro lado, é nosso desejo permitir que tal seja fácil de fazer, sem necessidade de conhecimentos especiais de programação. Assim, procurámos também desenvolver uma interface gráfica eficaz e usável para a extracção da informação relevante.

Finalmente, a informação apresentada será tão mais útil quanto mais fácil seja a sua compreensão. Estudámos pois qual a melhor forma de permitir a utilizadores comuns usar técnicas avançadas de visualização de informação para explorar e apresentar os dados recolhidos. Um protótipo, o Metabrain, foi criado para demonstrar a validade da nossa abordagem.

O artigo encontra-se dividido da seguinte forma: primeiro é feito um resumo dos trabalhos analisados mais importantes, seguido de uma descrição da solução proposta e seu desenvolvimento. Por fim são descritos os métodos de avaliação propostos.

## 2. TRABALHO RELACIONADO

As soluções analisadas partilham entre si a utilização da Internet e os serviços nela disponibilizados como fonte dos seus métodos de análise. Não existe a necessidade de catalogar ou estruturar manualmente informação nem são utilizadas bases de conhecimento especialmente criadas para o funcionamento destes trabalhos.

O trabalho *The Secret Lives of Numbers*<sup>1</sup> é possivelmente um dos primeiros trabalhos nesta área. Este permite visualizar a popularidade de um enorme conjunto de números inteiros num dado motor de pesquisa. Esta popularidade é estimada mediante a contagem dos resultados de uma simples pesquisa num motor de busca por cada número. A mesma técnica é em *Twitter Venn*<sup>2</sup> com o objectivo de analisar o relacionamento entre vários conceitos, com base nas redes sociais. Aqui, são registados o número de resultados onde várias combinações de palavras-chave co-ocorrem.

O projecto *Prism*<sup>3</sup> permite criar paletas de cores com base num dado conceito. A técnica utilizada é semelhante à anterior mas utiliza listas pré-definidas de cores e palavras relacionadas com saturação e luz. A paleta de cores é criada com base no número de resultados das várias combinações entre o conceito dado pelo utilizador e as listas existentes (maçã+verde+claro, etc.).

Recentemente foram tornados públicos os resultados de alguns projectos particulares como o *Simplistic Sentiment Mining from Tweets*<sup>4</sup> e [Kramer10] que utilizam APIs de redes sociais como base para uma análise de co-ocorrência de conceitos. O primeiro procura pelas palavras que mais ocorrem juntamente com o nome de alguns dispositivos comerciais. Este método simples produz resultados que permite, por exemplo, aos comerciantes verificar a popularidade dos seus produtos. O segundo trabalho analisa tudo o que é dito diariamente e faz uma contagem de palavras positivas vs palavras negativas, criando assim um gráfico de felicidade. Nos resultados, é fácil verificar que épocas como o Natal e a Páscoa, entre outras épocas festivas, correspondem a picos de felicidade.

Trabalhos como [Bollegala07], [Cimiano04] e [Etzioni04] utilizam o número de resultados devolvido pelos motores de pesquisa como fonte de informação para realizar inferências, através de por exemplo conectores linguísticos como os descritos em [Banko08]. Utilizando o conector “é um” e um conceito dado pelo utilizador, é possível realizar uma pesquisa, como por exemplo: “Portugal é um”, e verificar qual a palavra que mais ocorre de seguida. Desta forma a palavra que mais deve ocorrer será “país”, sendo assim possível descobrir de forma

automática a que categoria pertence um determinado conceito. O contrário também possível com a utilização do conector “tais como”. Com este conector é possível fazer uma pesquisa por “países tais como” e verificar quais as palavras que ocorrem de seguida nos resultados obtidos.

Todos estes trabalhos incluem restrições do domínio de execução, desenvolvimento fechado, restrições no método de visualização, impossibilidade por parte do utilizador de usar vários métodos em conjunto e por fim a necessidade, por parte do programador de desenvolver de raiz os métodos em questão.

## 3. SOLUÇÃO

Dado o nosso objectivo de permitir mesmo a utilizadores não experientes em programação a extracção e visualização de informação implícita na Web, desenvolvemos o *Metabrain* de forma modular e altamente personalizável, sem descurar a sua usabilidade e facilidade de utilização. Isto foi feito tanto para a extracção de informação como para a sua visualização, ao fornecer vários modos de visualização de informação os quais podem ser utilizados e personalizados de acordo com as necessidades do utilizador.

A arquitectura da nossa solução é composta por três camadas: biblioteca de extracção de informação; gestão de conjuntos de dados; visualização de dados, descritas em seguida.

### 3.1 Biblioteca Metabrain

A biblioteca *Metabrain* foi desenvolvida com o intuito de proporcionar um conjunto de métodos de análise e extracção de informação da Web prontos a usar de uma forma simples sem que os seus utilizadores tenham que se preocupar com a complexidade por detrás destes.

A biblioteca abstrai o seu utilizador do modo como a informação é obtida e pesquisada. Sendo incluídos na biblioteca métodos de acesso a várias APIs públicas, listas de palavras anotadas (por exemplo, palavras relacionadas com sentimentos) e técnicas de construção de *queries* complexas de forma automática.

Por exemplo, existem métodos que devolvem o número de resultados encontrados de várias palavras de uma lista, outros que analisam os resultados obtidos e conseguem identificar quais as palavras mais usadas e relacionadas com a nossa pesquisa inicial. Existem ainda outros que incluem técnicas especiais de geração de *queries*, os quais nos permitem extrair significado da Web. Um exemplo de uma das *queries* seria “Cidades tais como \*”, através da análise dos resultados obtidos é possível extrair da Web uma lista de cidades. Para além disto, estes métodos podem ser afinados de acordo com as necessidades do utilizador através de opções avançadas. Estas podem ser bastante úteis como por exemplo durante uma análise de sentimento. Seleccionada a opção de filtro de afirmações negativas durante a geração de *queries* de pesquisa estende a *query* normal ‘I like x’ para ‘I like x – “I don’t like”’, o que permite melhorar os resultados, excluindo alguns resultados negativos.

<sup>1</sup> <http://turbulence.org/Works/nums/> (visitado a 03/06/2010)

<sup>2</sup> <http://neoformix.com/Projects/TwitterVenn/view.php> (visitado a 03/06/2010)

<sup>3</sup> <http://nodebox.net/code/index.php/Prism> (visitado a 03/06/2010)

<sup>4</sup> <http://neoformix.com/2009/SimplisticSentimentMiningFromTweets.html> (visitado a 03/06/2010)

Tudo isto permite que o utilizador só tenha que preocupar com qual a informação que pretende extrair da Web e não em como essa extracção será realizada.

### 3.2 Aplicação Metabrain

A aplicação Metabrain é construída com base na biblioteca Metabrain. A interface gráfica desta aplicação foi desenvolvida para permitir que um comum utilizador consiga realizar pesquisas de informação simples ou avançadas e consiga analisar os resultados aí obtidos de uma forma fácil através de vários modos de visualização.

A utilização desta aplicação pode ser levada a cabo em 3 passos: gestão de conjuntos de dados (*datasets*); escolha de informação a visualizar e seu modo de visualização; e personalização da visualização e exportação desta.

#### 3.2.1 Conjuntos de dados

O primeiro passo na utilização da aplicação é a obtenção de informação implícita na Web. Para tal foi criada uma interface que dá ao utilizador uma grande liberdade na criação de novos conjuntos de dados.

O principal desafio no desenvolvimento desta interface foi permitir que o utilizador consiga executar diferentes tipos de queries de modo tão flexível quanto ao utilizar directamente a API da biblioteca Metabrain. Cada pesquisa tem como resultado uma tabela, com cada coluna correspondendo a uma faceta (número de resultados da pesquisa, por exemplo). Para além disto é possível encadear pesquisas, tornando possível, por exemplo, extrair uma lista de cidades da Web e de seguida encontrar quais as palavras que mais ocorrem junto a estas.

A solução encontrada para permitir o encadeamento de pesquisas foi possibilitar a realização de pesquisas hierárquicas. Durante uma pesquisa o utilizador pode decidir usar como parte da sua *query* um valor proveniente do resultado de uma outra utilizando uma variável que tomará sucessivamente os vários valores de uma das colunas do resultado da pesquisa original. A Figura 1 Exemplo de pesquisa de informação. Figura 1 exemplifica uma possível pesquisa hierárquica para exemplo dado anteriormente. Neste o utilizador decide obter uma lista de palavras relacionadas com cada elemento dos resultados da variável %1, que corresponde à extracção, da Web, de uma lista de cidades.

Como durante a criação de pesquisas encadeadas avançadas o utilizador pode necessitar de executar algum nó da pesquisa de forma a validar os seus resultados em separado, cada pesquisa da hierarquia é acompanhada na interface por um botão *preview*.

Este mecanismo de pesquisa foi desenvolvido com o auxílio de tecnologias DHTML (CSS, JavaScript) de modo a criar um ambiente de maior dinamismo e rapidez de utilização. A junção destes vários elementos cumpre o nosso objectivo de disponibilizar ao utilizador uma grande liberdade na criação de novos conjuntos de dados.

Tendo em conta que alguns dos métodos de análise necessitam de fazer, por vezes, milhares de pesquisas online e que a maior parte dos serviços online usados impõe limites de utilização, foi desenvolvida uma cache que permite utilizar qualquer um dos resultados obtidos

em pesquisas anteriores como fonte para novas pesquisas. Todos os resultados ficam também disponíveis para posterior visualização (**Error! Reference source not found.**).

Figura 1 Exemplo de pesquisa de informação.

ID	DATA SET	TYPE	PARAMS	ADDED	EDIT	DELETE	EXISTING VISUALIZATIONS	VISUALIZE
5	Sample Name	Search count	query = %1	2010-04-27 18:20:19.987000	Edit	Delete		New
7	Sample Name	Search suggests	query = obama	2010-04-27 22:12:15.874000	Edit	Delete		New
8	Sample Name	Search count	query = %1	2010-04-28 13:39:17.184000	Edit	Delete		New
9	Sample Name	Search count	query = %1	2010-05-09 15:52:46.729000	Edit	Delete		New
10	Sample Name	Map join	join1 = %1, join2 = %1	2010-05-09 21:44:39.701000	Edit	Delete		New

Figura 2 Listagem de conjuntos de dados guardados na base de dados.

#### 3.2.2 Visualização

O objectivo da secção de visualização é permitir que utilizador consiga visualizar a sua informação para que esta tenha um maior significado, mais uma vez dando um elevado grau de liberdade na sua personalização.

De forma a cumprir este objectivo a interface é dividida em duas secções: escolha dos dados a visualizar e modo de visualização; e pré-visualização mais personalização. Cada conjunto de dados, como já referido, tem a forma de uma tabela com várias colunas ou facetas.

Existem vários modos de visualização disponíveis, desde gráficos de linha até *treemaps*. Cada uma destas visualizações é caracterizada por um certo número de graus de liberdade, que definem em cada visualização quais os elementos desta que podem variar. Depois de analisado um grande número de visualizações concluímos que estes graus de liberdade podem ser encapsulados num destes quatro campos: colunas (eixo X), linhas (eixo Y), cor/paleta de cores, tamanho (largura/altura). Os dois primeiros campos são usados na selecção dos dados a visualizar e os outros dois são usados para personalizar a visualização.

Tendo isto em conta e sabendo que o utilizador ao criar uma visualização já tem em mente quais os dados que pretende usar, mas nem sempre sabe qual a melhor forma



de os visualizar, decidimos criar uma interface que o auxilie nesta tarefa. A interface foi então desenvolvida de modo a que à medida que o utilizador selecciona as facetas que pretende visualizar uma lista de sugestões seja dinamicamente actualizada.

As facetas que o utilizador pode seleccionar são recuperadas do conjunto de dados seleccionado e listadas na interface. De forma a facilitar a selecção destas facetas, a sua listagem é automaticamente separada em categorias, de acordo com o conteúdo de cada uma. Os elementos desta lista podem ser arrastados pelo utilizador para os campos que representam os eixos X e Y, indicando que devem ser usados na visualização.

De acordo com as facetas seleccionadas e sua localização nos eixos, todas as visualizações incluídas na aplicação são verificadas. No fim desta verificação, todas as visualizações cujo número de graus de liberdade seja compatível com a selecção do utilizador será apresentada a este como uma sugestão de visualização. O utilizador poderá então seleccionar qual a visualização que pretende usar.

Depois de seleccionada uma visualização, esta é apresentada automaticamente utilizando os dados seleccionados. Nesta altura são também adicionados novos campos na interface que permitem uma maior personalização da visualização seleccionada. Aqui o utilizador poderá decidir, no caso de um gráfico de barras por exemplo, que o tamanho das barras e/ou sua cor é definido pelo valor de outra coluna, simplesmente arrastando essa coluna da lista anteriormente referida para o campo de personalização em questão.

Os métodos de visualização usados nesta aplicação são desenvolvidos utilizando o *toolkit* Protovis [Bostock09] que proporciona uma grande flexibilidade na criação de visualizações em Javascript, permitindo assim a exportação de visualizações criadas para outros websites.

Na Figura 3 é apresentado um exemplo de visualização através de um gráfico de barras.

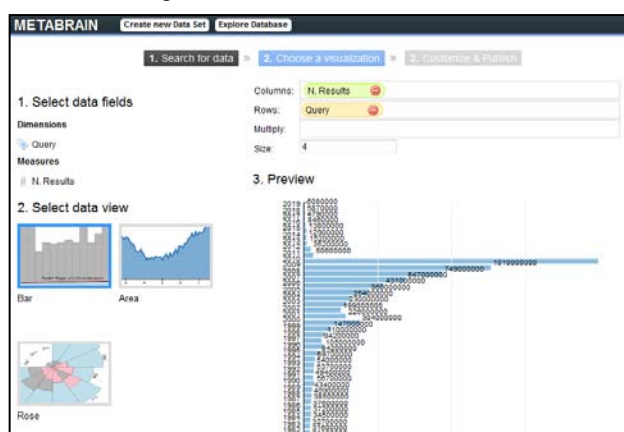


Figura 3 Exemplo de visualização de um conjunto de dados.

### 3.3 Futuro

Como trabalho futuro, está em primeiro lugar a finalização da aplicação, no que toca à personalização das visualizações, seguida do desenvolvimento de novos módulos de análise e novos modos de visualização.

## 4. AVALIAÇÃO

Apesar de estarmos confiantes na qualidade da solução desenvolvida, esta será avaliada oportunamente, para demonstrar inequivocamente a sua adequação ao problema em causa. Em primeiro lugar, faremos uma avaliação heurística para descobrir os principais problemas de usabilidade da interface. Uma vez corrigidos esses problemas, terão lugar testes com utilizadores em que estes deverão desempenhar um conjunto de tarefas de extracção e visualização de informação. Serão medidos tempo e erros, e avaliada a sua satisfação e a facilidade de aprendizagem e compreensão da interface.

## 5. CONCLUSÃO

Durante o desenvolvimento deste trabalho foram analisados vários trabalhos que utilizam métodos estatísticos, análise de padrões e tendências como meio para aceder ao “inconsciente colectivo” que é a Internet. Desta análise foi possível descobrir os vários métodos existentes actualmente e quais as melhores fontes na Web onde estes podem ser usados. As várias abordagens analisadas, apesar de se basearem na mesma ideologia, apresentam alguns problemas e nenhuma tira proveito de todos os métodos de extracção analisados. Para além disso, também restringem o domínio de execução e apresentam pouca flexibilidade na visualização dos resultados.

De modo a responder aos problemas encontrados surge a necessidade da criação de uma solução que disponibilize ao utilizador uma vasta quantidade de métodos de extracção de informação, sem que este se tenha de preocupar com análise de linguagem natural ou criação de bases de dados gigantes de conhecimento anotado. A solução proposta materializa-se sob a forma da biblioteca Metabrain e uma aplicação complementar que fornece ao utilizador um modo personalizável de obter e visualizar a informação a qual, de outro modo, seria difícil obter e analisar.

## 6. REFERÊNCIAS

- [Bollegala07] Bollegala, Danushka, Yutaka Matsuo, e Mitsuru Ishizuka. “Measuring semantic similarity between words using web search engines.” *ACM*, 2007. 757--766.
- [Banko08] Banko, Michele, e Oren Etzioni. “The Tradeoffs Between Open and Traditional Relation Extraction.” *Association for Computational Linguistics*, 2008. 28--36.
- [Bostock09] Bostock, Michael, e Jeffrey Heer. “Protovis: A Graphical Toolkit for Visualization.” *IEEE Transactions on Visualization and Computer Graphics* (IEEE Educational Activities Department) 15 (2009): 1121--1128.
- [Cimiano04] Cimiano, Philipp, e Steffen Staab. “Learning by googling.” *SIGKDD Explor. Newsl.* (ACM) 6 (2004): 24--33.
- [Etzioni04] Etzioni, Oren, et al. “Web-scale information extraction in knowitall: (preliminary results).” *ACM*, 2004. 100--110.
- [Kramer10] Kramer, Adam D. “An unobtrusive behavioral model of gross national happiness.” *ACM*, 2010. 287--290.



# Pontos de Interesse Relevantes num Mapa: à Procura de Boas Práticas

Filipe Gil      Ana Paula Cláudio      Maria Beatriz Carmo  
Departamento de Informática, Faculdade de Ciências da Universidade de Lisboa  
fgil@armail.pt {apc,bc}@di.fc.ul.pt

## RESUMO

*Neste artigo tentamos identificar um conjunto de boas práticas para a produção de símbolos gráficos que representam pontos de interesse relevantes sobre um mapa, quando se usam computadores de secretária ou portáteis. Recorrendo a testes com utilizadores, analisámos a eficácia de um conjunto de símbolos gráficos criteriosamente escolhidos. O protótipo interactivo implementado, no contexto do qual os testes tiveram lugar, incorpora um mecanismo de filtragem que determina os pontos de interesse que são efectivamente relevantes para o utilizador e calcula os correspondentes níveis de relevância; a cada nível de relevância corresponde um símbolo gráfico específico.*

## Palavras-chave

*Pontos de interesse, relevância, símbolos gráficos, variáveis visuais.*

## 1. INTRODUÇÃO

A utilização da Internet, como sabemos, cresceu exponencialmente a partir do início da década de 1990. No que respeita à utilização de mapas na Internet podem ser identificadas três etapas. Na primeira, os mapas existentes em papel foram digitalizados e disponibilizados em imagens estáticas, tal como os seus congéneres originais em papel. Na segunda etapa, após 1997, a Web passou a disponibilizar mapas interactivos, permitindo que acções do utilizador conduzam a alterações do seu aspecto em tempo real. Mais recentemente a investigação centra-se em aspectos específicos deste novo tipo de mapas, no seu desenho, no modo como são usados, no modo como se disponibilizam.

Temos vindo a trabalhar num projecto que explora a visualização de pontos de interesse sobre um mapa, quer em dispositivos móveis [Carmo08], quer em computadores de secretária ou portáteis [Paiva09]. O trabalho apresentado neste artigo insere-se nesta segunda vertente. Um dos nossos objectivos é identificar símbolos gráficos que transmitam a relevância de forma adequada ao utilizador, tendo em consideração que um estímulo bem escolhido pode reforçar o sinal produzido por esse estímulo, de forma a ser processado mais rapidamente [Wolfe04].

No protótipo que desenvolvemos, cada símbolo gráfico desenhado sobre o mapa: (i) revela a localização geográfica de um ponto de interesse, (ii) indica visualmente o tipo de item representado e (iii) mostra a relevância que esse ponto tem para o utilizador. Esta relevância é calculada por uma função matemática, tendo

por base informação que o utilizador fornece à aplicação de forma interactiva. Cada símbolo possui duas características distintas: o aspecto gráfico (tamanho, forma, cor, entre outros) e o conteúdo, que se refere à semântica que lhe associamos (o símbolo representa um hotel, uma igreja, por exemplo). Ambas as características têm de ser tratadas de forma sistemática e cuidadosa. Foi esta a abordagem que seguimos no estudo apresentado.

Na secção 2 descrevemos sucintamente um conjunto de trabalhos que se debruçam sobre os atributos visuais que despertam a atenção humana. A secção 3 é dedicada ao protótipo implementado e aos testes efectuados com utilizadores. Na secção 4 é feita a discussão dos resultados e apontam-se linhas de trabalho futuro.

## 2. OS ATRIBUTOS VISUAIS QUE DESPERTAM A NOSSA ATENÇÃO

O cartógrafo francês Jacques Bertin foi o primeiro a propor uma abordagem sistemática para a comunicação de informação na forma gráfica no contexto particular dos mapas, na sua obra *Sémiologie Graphique* de 1967, traduzida para língua inglesa em 1983 [Bertin83]. Além das duas variáveis,  $x$  e  $y$ , que fornecem a posição no plano de um mapa, Bertin identificou seis características das marcas ou símbolos gráficos, que designou por *retinal variables*: o tamanho, a textura, a orientação, a forma e mais duas relativas à cor, a cor dominante (*color hue*) e o valor ou brilho (*color value*). Na opinião deste autor, o tamanho está no topo da lista, uma vez que é uma variável quantitativa, ordenada e associativa (i.e., permite identificar os objectos pertencentes a um grupo), enquanto

a orientação se encontra no lugar mais baixo da mesma lista, tendo apenas a característica associativa.

O estudo de Bertin [Bertin83] não contém nenhuma referência a trabalho prévio que apoie as ideias nele apresentadas. Contudo, este seu trabalho constitui um marco que é amplamente citado e cujos conceitos foram comprovados, em grande medida, por estudos efectuados posteriormente.

De entre os trabalhos mais recentes sobre atributos visuais no contexto particular da geovisualização, destacamos [Swienty08] e [Garlandini09], ambos usando, entre outros, resultados obtidos por Wolfe e Horowitz. Estes dois últimos autores, professores de Oftalmologia da Harvard Medical School, fizeram um estudo exaustivo sobre os atributos gráficos que despertam a nossa atenção, sem particularizarem nenhum contexto aplicacional [Wolfe04]. Baseando-se num conjunto significativo de estudos de diversos autores, principalmente das áreas das neurociências e da psicologia, Wolfe e Horowitz definem cinco grupos de atributos visuais baseados na probabilidade que estes têm de despertar a atenção do observador.

Swienty et al. apresentam uma tabela de atributos visuais [Swienty08] que agrega as propostas de Bertin e de outros autores como MacEachren [MacEachren95]. Desta tabela destacamos o movimento que aparece decomposto em diversas vertentes. Swienty et al. propõem uma metodologia para produzir geovisualizações que envolve o cálculo da relevância com base no contexto de uso. Os objectos são filtrados em função desse contexto e a hierarquia dos valores da relevância é representada recorrendo a atributos visuais. Os autores apresentam exemplos ilustrativos de aplicação da metodologia, usando como método de pré-avaliação o registo e posterior análise dos movimentos dos olhos dos observadores.

Os estudos de Garlandini e Fabrikant vieram também confirmar algumas das ideias de Bertin [Garlandini09]; estas autoras testaram quatro variáveis visuais: tamanho, orientação, cor dominante e brilho. Estudando os movimentos dos olhos dos observadores e centrando a sua avaliação na tarefa de detectar a informação quando um estímulo visual ocorre, concluíram que a orientação produzia os piores resultados e que a variável tamanho é aquela com a qual se obtêm observações mais eficazes e mais eficientes.

### 3. O ESTUDO REALIZADO

#### 3.1 O Protótipo Viswide

O protótipo utilizado neste estudo designa-se por VisWide e é destinado à visualização de pontos de interesse sobre mapas em computadores de secretária ou portáteis. A sua versão inicial encontra-se descrita em [Paiva09].

A construção do VisWide recorre a JavaScript, PHP, Java, SQL, HTML e SVG para o desenho dos mapas e dos símbolos gráficos que representam os pontos de interesse. Os atributos gráficos dos símbolos variam de modo a transmitir visualmente o nível de relevância do ponto de interesse que representam (Fig. 1).

A relevância que cada ponto de interesse tem para o utilizador é um valor numérico entre 0 e 1, calculado por uma função de grau de interesse [Carmo08]. Com base na informação fornecida pelo utilizador, através de uma interface específica, esta função calcula o valor que expressa a relevância de cada item e apenas aqueles para os quais este valor está acima de um limiar, definido pelo utilizador, são representados sobre o mapa.

Na versão actual do protótipo VisWide, a relevância é transformada num valor discreto dividindo o intervalo entre o limiar e o valor 1 em 3 subintervalos iguais, correspondendo cada um deles a um nível de relevância: pouco relevante, relevante e muito relevante. Este valor discreto irá determinar a escolha da simbologia.

Optámos por usar três subintervalos, considerando que este número é indicado para, por um lado, dar informação suficiente ao utilizador (consideramos que usar apenas dois intervalos é insuficiente) e, por outro lado, não o sobrecarregar com uma carga cognitiva excessiva. Esta mesma abordagem é usada por Swienty et al. [Swienty08].

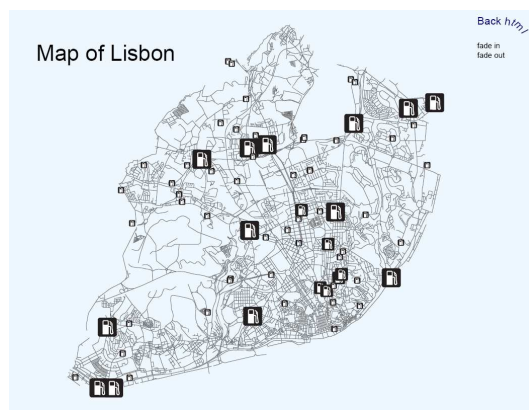


Fig. 1- Aplicação VisWide mostrando o mapa de Lisboa com os pontos de interesse nele desenhados

#### 3.2 Testes de usabilidade

##### 3.2.1 Realização dos testes

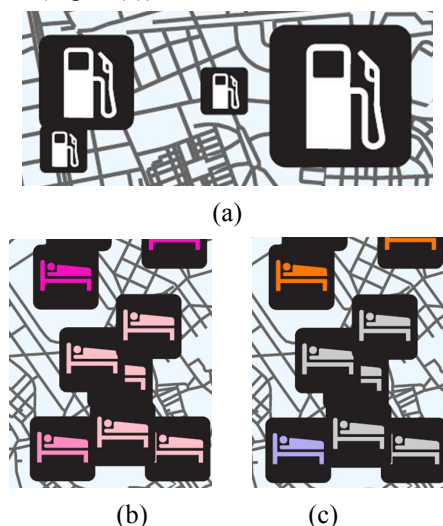
Foram entrevistados 16 utilizadores divididos de forma equilibrada em três faixas etárias (20-30, 30-45, +45), de formação base variada (alunos universitários de Informática, Biologia, Arquitectura e Design, professores do ensino secundário e universitário de diversas áreas, advogados), de ambos os géneros (50% de cada género). Todos os utilizadores usam mapas na Internet, metade deles diariamente, sem predominância de faixa etária. Um dos utilizadores, na faixa etária 20-30, é daltónico.

O mapa utilizado foi sempre o da cidade de Lisboa, que todos os utilizadores conhecem bem. Foram testadas diferentes simbologias no protótipo, escolhidas tendo em conta: 1) os estudos efectuados pelos autores que referimos anteriormente, e 2) uma pré-avaliação que deixou de fora as hipóteses consideradas menos interessantes.

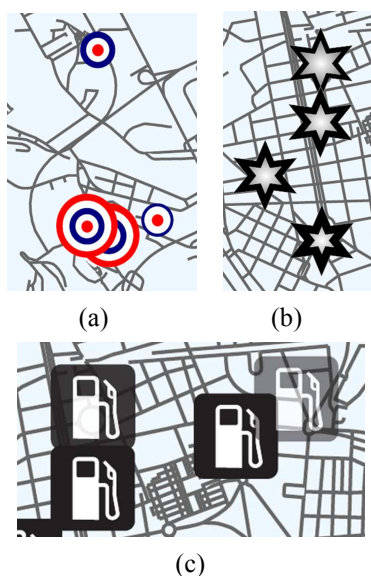
Designamos por versão cada conjunto de simbologias testadas. O nosso estudo envolveu cinco versões, variando a forma dos símbolos e/ou os seus atributos visuais.

Na versão 1 utilizámos símbolos com semântica óbvia.

Neste caso a relevância é transmitida pelo tamanho do símbolo Fig. 2(a). As versões 2, 3 e 5 recorrem aos mesmos símbolos base. Para expressar a relevância, a versão 2 usa três níveis distintos de saturação (Fig. 2(b)), a versão 3 usa três cores distintas (Fig. 2(c)) e a versão 5 usa movimento (Fig. 3 (c)).



**Fig. 2:** (a) versão 1: o símbolo com maior tamanho é o mais relevante; (b) versão 2: o símbolo cor de rosa mais saturado é o mais relevante; (c) versão 3: o símbolo laranja é o mais relevante; o lilás é o intermédio e o cinzento é o menos relevante.



**Fig. 3:** (a) versão 4- alvo com mais anéis é o mais relevante (b) versão 4- estrela com rebordo mais grosso é a mais relevante ; (c) versão 5- símbolo que pisca mais rapidamente é o mais relevante.

Na versão 4 os símbolos utilizados não têm semântica óbvia associada. Testámos apenas dois símbolos, para duas categorias de pontos de interesse: uma estrela e um alvo. O alvo com mais anéis corresponde ao nível mais elevado de relevância; os alvos relativos à menor relevância e à relevância intermédia têm o mesmo número de anéis, mas a espessura do anel exterior é menor no primeiro (Fig.

3(a)). Na estrela, uma maior espessura do rebordo corresponde a uma maior relevância (Fig. 3(b)).

A aplicação do atributo cor, usado nas versões 2 e 3 para representar o nível de relevância, tem como desvantagem óbvia o facto de existir uma fatia da população que sofre de deficiências ao nível da percepção da cor [Ware00]. Adicionalmente, o efeito da cor perde-se em grande medida quando se visualiza a preto e branco (por exemplo, quando se imprime numa impressora a preto e branco).

Cada utilizador realizou os testes em três etapas:

Etapa 1 - Pesquisas Controladas: o utilizador de teste analisou os *layouts* resultantes de um conjunto de pesquisas fixas, predefinidas, de forma a garantir que em todas as pesquisas aparecessem símbolos relativos aos três níveis de relevância. Para cada *layout* respondeu a um conjunto de perguntas.

Etapa 2 - Pesquisas Livres: o utilizador de teste analisou os *layouts* resultantes de um conjunto de pesquisas efectuadas livremente por ele próprio, podendo confirmar ou ajustar respostas dadas na etapa anterior.

Etapa 3- Balanço geral: o utilizador ordenou as três versões que considerou mais adequadas. Pôde ainda dar sugestões de melhoramento e opiniões pessoais de carácter mais genérico sobre o tipo de representações em estudo.

As perguntas feitas aos utilizadores foram as seguintes:

1. Quantos ícones diferentes conseguem distinguir-se sobre o mapa?
2. Como associa cada ícone ao seu nível de relevância?
3. Qual foi o ícone que lhe chamou a atenção em primeiro lugar?
4. Tendo em conta que se pretende avaliar com estes testes a qualidade dos ícones usados para expressar o Valor de Relevância de cada Ponto de Interesse, como classificaria cada versão da aplicação (Muito Boa, Média, Fraca)?
5. Qual o ícone mais sugestivo, de todos os que observou?
6. Quais são, por ordem, as três versões que considera mais adequadas?

### 3.2.2 Resultados e Discussão

Analisemos as respostas dos utilizadores a cada uma das perguntas referidas na subsecção anterior:

1. Todos os utilizadores conseguiram distinguir três símbolos diferentes sobre o mapa em todas as versões, mesmo o utilizador de teste daltónico.
2. A versão 1 foi a única em que todos os utilizadores acertaram no nível de relevância dos símbolos gráficos; a que apresentou piores resultados foi a versão 5 com apenas 9 respostas 100% correctas. As restantes 3 versões apresentam resultados muito análogos, apenas um ou dois utilizadores se confundiram ou mostraram hesitação na resposta.
3. Também neste caso, a versão 1, recebeu respostas correctas de todos os utilizadores. Para todas as outras versões, a atenção de 11 ou 12 utilizadores foi despertada em primeiro lugar pelo símbolo que

representava o nível de relevância mais elevado.

- O gráfico da Fig. 4 mostra as médias das classificações obtidas para as cinco versões que testámos. A versão 1 foi a que apresentou uma melhor média na classificação, 2,7 com um valor de desvio padrão de 0,6. A versão 5 teve a menor classificação média, 1,8 com desvio padrão de 0,66, seguida de perto pela versão 3.

Todas as versões receberam pelo menos uma vez a classificação máxima e a mínima. A Moda e a Mediana da versão 1 têm o valor 3 e nas restantes versões tem o valor 2. A versão 4 foi aquela que levantou mais hesitações nos utilizadores quanto à classificação.

- O símbolo que recebeu mais impressões favoráveis, 11, foi o Mais Relevante da versão 1, seguido pelos símbolos de Mais Relevantes da versão 3 e da versão 4, com 5 respostas. A versão 2 foi a que teve os símbolos menos votados, apenas 2 votos para o símbolo Mais Relevante.
- O gráfico da Fig. 5 mostra o *ranking* das 5 versões. A versão 1 foi a melhor classificada, seguida pela versão 4 que foi a que ficou mais vezes em segundo e em terceiro lugares. A versão 5 destaca-se pela negativa.

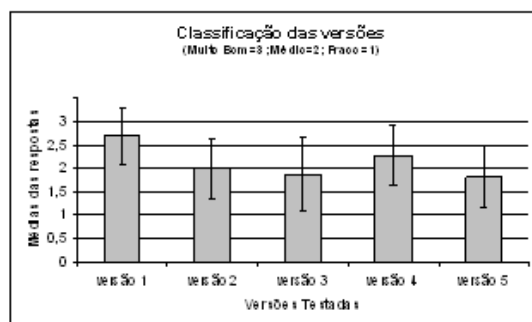


Fig. 4- Gráfico que ilustra as classificações obtidas para as cinco versões testadas.

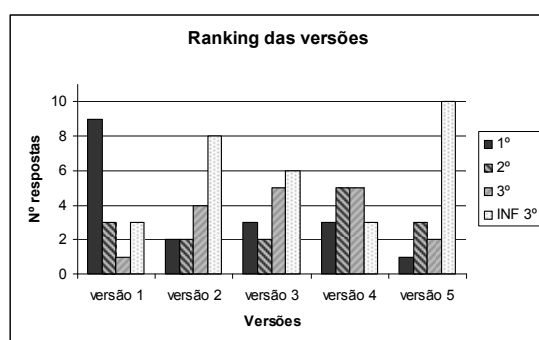


Fig. 5- Gráfico que mostra o ranking das cinco versões testadas (INF 3º designa as classificações inferiores ao 3º lugar).

Constatou-se que a versão que mais confundiu os utilizadores, no que diz respeito ao nível de relevância, foi a versão 5 em que se usou o atributo movimento. Este resultado é contrário à classificação de [Wolfe04] que incluía o movimento no grupo dos atributos prováveis.

Os resultados francamente positivos obtidos pela versão 1, na qual se aplicou o atributo visual tamanho para

representar a relevância, confirmam as ideias base de [Bertin83] e de [Wolfe04]. Esta versão, que usa símbolos com semântica óbvia associada, é também a que recebeu mais classificações de Muito Bom (Fig. 4) e é a preferida dos utilizadores (Fig. 5).

De um modo geral, nas respostas dadas pelos utilizadores, não se observaram diferenças significativas, quer entre faixas etárias distintas, quer entre géneros. O utilizador daltónico não teve qualquer dificuldade em realizar os testes que envolviam cor. As cores que conseguiu visualizar mantiveram a semântica no que diz respeito aos níveis de relevância.

#### 4. TRABALHO FUTURO

No futuro próximo vamos prosseguir os nossos testes, incluindo mais utilizadores daltónicos. Iremos reflectir sobre várias sugestões que nos foram dadas pelos utilizadores de teste e, eventualmente, proceder ao teste de outras versões. Numa etapa seguinte pretendemos incorporar estes resultados com tratamento de *cluttering* e pensar como poderemos aplicar os resultados obtidos a outro tipo de dispositivos com características físicas e contextos de utilização diferentes.

#### 5. AGRADECIMENTOS

Agradecemos à FCT e ao LabMAg o apoio financeiro, ao Bruno Paiva a concepção e a implementação da versão inicial da aplicação, à Profª Drª Cristina Catita o mapa da cidade de Lisboa em formato vectorial. Para finalizar, agradecemos a colaboração de todos os utilizadores de teste.

#### 6. REFERÊNCIAS

- [Bertin83] J. Bertin. *Semiology of Graphics: Diagrams, Networks, Maps*. University of Wisconsin Press, 1983
- [Carmo08] M. B. Carmo, A. P. Afonso, P. P. Matos, A. Vaz. *MoViSys - A Visualization System for Geo-Referenced Information on Mobile Devices*. VISUAL 2008, pp 167-178, 2008
- [Garlandini09] S. Garlandini, S. I. Fabrikant. Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization. *Proceedings COSIT 2009*, pp 195-211, LNCS 5756/2009
- [MacEachren95] A. M. MacEachren. *How Maps Work. Representation, Visualization and Design*. Guilford Press, 1995
- [Paiva09] B. Paiva, A. P. Cláudio, M. B. Carmo, P. Pombinho, C. Catita. *Estratégias para a Representação de Pontos de Interesse sobre Mapas*. Actas do 17º EPCG, pp 305-306, 2009
- [Swienty08] O. Swienty, T. Reichenbacher, S. Reppermund, J. Zihl. The role of relevance and cognition in attention-guiding geovisualisation. *The Cartographic Journal*, Vol.45 (3), pp. 227-238, 2008
- [Ware00] C. Ware. *Information Visualization - Perception for Design*, Morgan Kaufmann, 2000
- [Wolfe04] J. M. Wolfe, T. S. Horowitz. Opinion: What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews. Neuroscience* 5(6) pp.495-501, 2004

# Pesquisa Interactiva de Fotografias com Base na Regra dos Terços

Gabriel Barata Ricardo Dias Sandra Gama Manuel J. Fonseca Daniel Gonçalves

Departamento de Engenharia Informática  
INESC-ID / IST / Universidade Técnica de Lisboa  
R. Alves Redol, 9, 1000-029 Lisboa, Portugal

[gabriel.barata, ricardo.dias, sandra.gama]@ist.utl.pt, [mjf, daniel.goncalves]@inesc-id.pt

---

## Resumo

*Hoje em dia, geramos uma quantidade cada vez maior de fotografias digitais. Estas encontram-se, frequentemente, desorganizadas ou organizadas em hierarquias complexas. A pesquisa por determinadas imagens específicas torna-se, assim, difícil. Neste artigo apresentamos uma solução para a pesquisa interactiva de fotografias que combina técnicas de visualização interactiva com algoritmos de recuperação de imagem. Para a visualização utilizamos histogramas multifacetados e introduzimos uma nova técnica de apresentação dos resultados chamada Linha de Similaridade. Adicionalmente, comparamos duas formas de descrever o conteúdo das fotografias usando histogramas locais; uma que divide a fotografia em nove regiões iguais e outra que usa a regra dos terços para determinar as regiões. Finalmente, incluímos também um algoritmo para contar o número de faces presentes nas fotografias. Testes preliminares mostram que a utilização da regra dos terços para calcular os histogramas locais apresenta resultados comparáveis com a solução tradicional, com a vantagem de usar menos histogramas.*

## Palavras-Chave

*Recuperação de Fotografias, Visualização, Interface Utilizador, Interação Pessoa-Máquina*

---

## 1. INTRODUÇÃO

As câmaras digitais estão presentes no nosso quotidiano, sob a forma de máquinas fotográficas digitais comuns e de telemóveis. Tiramos uma quantidade cada vez maior de fotografias, descarregando-as para o sistema de ficheiros, muitas vezes sem grande preocupação quanto à organização. Como consequência, o processo de encontrar uma fotografia específica é cada vez mais complexo, o que faz com que o desafio de recuperar fotografias digitais tenha vindo a ganhar uma relevância significativa.

De facto, o armazenamento utilizando um sistema de ficheiros implica uma classificação hierárquica que, além de exigir um elevado esforço para manter alguma organização, não tira partido da informação contida nas próprias imagens [Lee 09]. Além disso, a pobre interactividade que este método apresenta leva o utilizador a um elevado esforço para recuperar as suas fotografias.

A extracção de características relativas às imagens (como texturas, cores dominantes e formas, entre outros), permite aplicar estratégias de classificação aos dados [Stojanovic 07] e promover a recuperação de imagens baseada no seu conteúdo, potencializando o desenvolvimento de ferramentas interactivas para pesquisa. Ferramentas como o VisualSEEK [Smith 97], que se baseia nas cores dominantes e usa histogramas HSV, têm sido usadas

como ponto de partida para outras investigações. Esta aplicação permite ao utilizador desenhar um conjunto de regiões (em que especifica a posição, a dimensão e a cor), devolvendo as imagens que melhor correspondem ao desenho. Contudo, além da cor, pode ser interessante considerar múltiplos parâmetros para pesquisa interactiva de fotografias. A pesquisa facetada, que dá liberdade aos utilizadores para explorar relações na informação, tem vindo a ser explorada nos últimos tempos.

Um exemplo desse tipo de abordagem é o FacetLens [Lee 09], um sistema de visualização interactiva, que suporta a exploração e compreensão de informação, utilizando conjuntos de dados facetados.

A solução que apresentamos aplica este conceito à imagem, fornecendo uma aplicação para pesquisa interactiva de fotografias. A interface que criámos baseia-se em duas técnicas principais de visualização que se complementam mutuamente: histogramas com diversas facetas e linhas de similaridade. A informação relativa aos resultados da pesquisa efectuada pelo utilizador é apresentada de forma completa, mas ao mesmo tempo simples e directa.

A abordagem que seguimos para a recuperação de fotografias baseia-se na criação de histogramas locais da imagem e na extracção de rostos humanos. Adicionalmente, usamos a Regra dos Terços para especificar as regiões da imagem



utilizadas para construir os histogramas locais.

Na secção 2 apresentamos os algoritmos usados para descrever o conteúdo e recuperar fotografias. Na secção 3 descrevemos a interface utilizador. A secção 4 refere-se aos resultados dos testes relativos à recuperação de fotografias, e finalmente, na secção 5 apresentamos as conclusões e trabalho futuro a realizar.

## 2. ALGORITMOS DE RECUPERAÇÃO

A nossa aplicação para recuperação interactiva de fotografias baseia-se na criação de histogramas locais da imagem e na extracção de rostos.

### 2.1. Histogramas Locais

Wyszecki e Stiles [Wyszecki 82] investigaram a percepção de cor pelo olho humano e concluíram que o nosso sistema visual consegue distinguir 13 cores básicas: vermelho, laranja, amarelo, verde, azul-esverdeado, azul claro, azul, roxo, rosa, castanho, branco, cinzento e preto. Tal como Belongie et al. [Belongie 01], utilizamos esta teoria.

A primeira abordagem para construção de histogramas locais consistiu em dividir a imagem em nove porções de igual dimensão. Numa segunda aproximação, decidimos adoptar a regra dos terços [Rul10], que tem sido amplamente usada em artes visuais para criar imagens equilibradas e com maior impacto. Tal como ilustrado na Figura 1, esta regra consiste na divisão da fotografia em nove partes iguais, usando duas linhas horizontais e duas linhas verticais igualmente espaçadas, de forma a que os elementos de interesse fiquem posicionados sobre as linhas e nas intersecções entre estas.



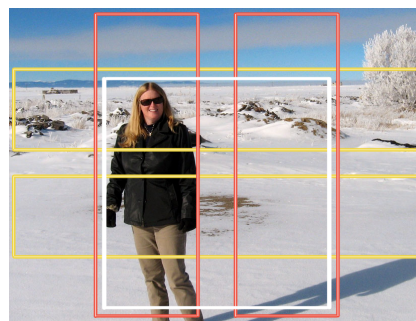
**Figura 1. Regra dos terços**

Considerámos cinco regiões para a construção de histogramas: as quatro áreas sobre as linhas que subdividem a imagem e uma região central (para maior tolerância relativamente a fotografias que não sigam a regra dos terços), como ilustrado na Figura 2.

### 2.2. Número de Faces

Além da criação de histogramas, é considerado o número de rostos numa imagem como parâmetro para recuperação de fotografias.

A extracção do número de faces da imagem é feita através de um classificador de Haar em cascata, que se baseia na



**Figura 2. Regiões de interesse: áreas em redor das linhas de separação utilizando a regra dos terços e zona central da imagem.**

investigação de Viola e Jones [Viola 01]. Este método consiste em distinguir padrões faciais de padrões não faciais através da subdivisão de uma imagem em janelas de menor dimensão e em submeter cada uma das regiões resultantes a um conjunto de testes em cascata em que serão aceites caso correspondam a padrões faciais e rejeitadas caso contrário.

### 2.3. Indexação

Dois mecanismos de indexação são empregues: um mecanismo de indexação multidimensional, de nome NBTree [Fonseca 03], que permite efectuar pesquisas pelos  $k$  vizinhos mais próximos de forma eficiente, e o tradicional índice invertido.

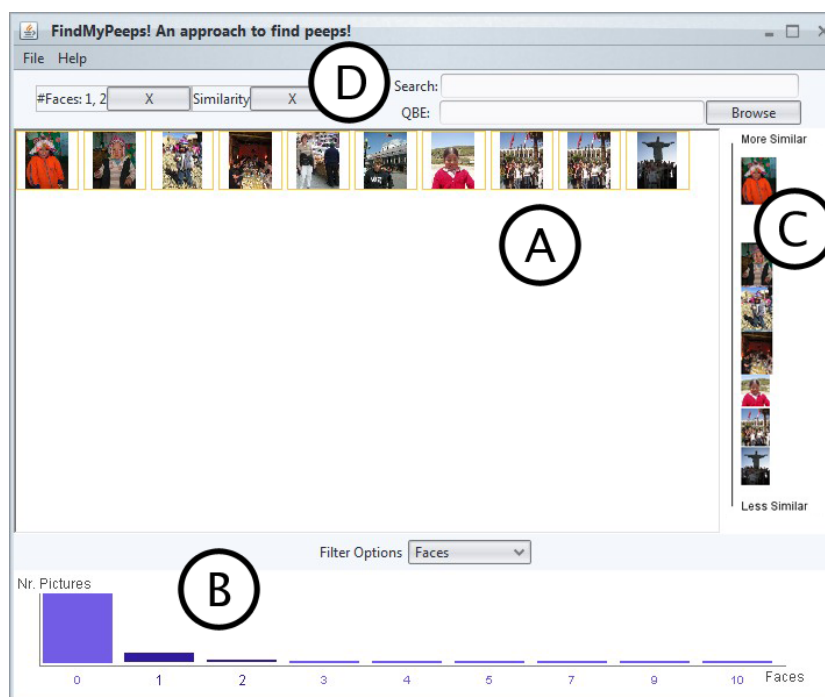
Na estrutura de indexação multidimensional guardamos os descritores de dimensão 14, obtidos pela concatenação do histograma (de dimensão 13, uma para cada cor) e do número de faces. Os descritores das várias regiões da fotografia são guardados na mesma estrutura de indexação.

Os dados referentes à data, orientação, nome, extensão, etc., são guardados em índices invertidos (um para cada tipo de dados).

Quanto às interrogações, o sistema suporta interrogações por exemplo, ou usando as características específicas. Na pesquisa por exemplo, o utilizador fornece uma imagem, a partir da qual se extrai um descritor que é usado para realizar uma interrogação KNN (*K Nearest Neighbours*) na estrutura multidimensional. Os descritores devolvidos correspondem às regiões mais semelhantes. Quantas mais regiões uma imagem tiver na lista devolvida, mais semelhante é a interrogação.

No caso dos outros dados, o utilizador usa o valor de uma propriedade para interrogar o índice invertido correspondente.

É de salientar que as interrogações podem ser armazenadas e combinadas de forma a gerar outras mais complexas. Por exemplo, podemos procurar todas as fotografias semelhantes a uma dada paisagem de montanha, tiradas em 2005, cuja orientação é a de paisagem e que mostre três pessoas.



**Figura 3. Interface Utilizador, mostrando o histograma multifacetado (B) e a linha de similaridade (C).**

### 3. INTERACÇÃO E VISUALIZAÇÃO

#### 3.1. Interface Utilizador

A interface que propomos consiste num mecanismo interactivo para recuperação de fotografias baseado em duas técnicas de visualização principais: histogramas multifacetados e linha de similaridade.

A interface está subdividida em quatro áreas principais, como ilustrado na Figura 3.

Na zona (A) são representadas, inicialmente, todas as imagens disponíveis. Quando é feita uma pesquisa, é nesta área que os resultados são apresentados. Esta interface, simples e natural, é semelhante aos esquemas de exploração de informação existentes nas interfaces gráficas dos sistemas operativos convencionais.

No entanto, quando uma fotografia é seleccionada na zona de navegação (A), sofre uma ampliação, inspirada na técnica *fish-eye view*, apresentando dois botões, que permitem, respectivamente, procurar fotografias semelhantes (interrogação através de exemplo) e abrir a fotografia.

Na base desta área existe uma ComboBox que permite aplicar diferentes filtros (número de rostos, orientação, ano, extensão do ficheiro e palavras-chave), permitindo representar diferentes facetas relativas às imagens.

A região (B) consiste numa das técnicas de visualização que adoptámos: histogramas multifacetados. Quando um filtro é seleccionado, na ComboBox da área (A), referida acima, o histograma é actualizado de acordo com o parâmetro correspondente. Por exemplo, se seleccionar-

mos o filtro "Ano", o histograma passará a mostrar essa faceta, isto é, a distribuição de fotografias por ano.

No painel (C) é representada a segunda técnica de visualização utilizada: linha de similaridade, permitindo a consulta simples e rápida do grau de semelhança entre a fotografia de interrogação e os resultados.

Finalmente, a área (D), no topo, permite uma pesquisa textual convencional, procurando fotografias que contenham, no seu nome, as palavras-chave introduzidas na caixa de texto. Além disso, a interface possibilita ao utilizador a selecção de uma qualquer fotografia existente no seu sistema de ficheiros para uma interrogação por exemplo.

#### 3.2. Técnicas de Visualização

Os histogramas multifacetados permitem visualizar informação relativa aos vários parâmetros disponíveis na filtragem, ou seja, ano, número de rostos, extensão do ficheiro e orientação (retrato ou paisagem). Optámos por manter o histograma bastante simples, contendo apenas informação essencial. No entanto, quando o rato é posicionado sobre uma barra do histograma, o valor correspondente é mostrado. A selecção de múltiplos valores relativos a uma faceta, que correspondem às várias barras do histograma, é suportada, quer para valores consecutivos, quer para valores alternados.

A linha de similaridade consiste num ordenamento e distribuição dos resultados de uma interrogação, que permite aos utilizadores ter uma perspectiva visual da semelhança entre a interrogação e as fotografias recupe-



radas, e destas entre si (Figura 3, área (C)).

Para tal, usamos a medida de semelhança entre a interrogação e os resultados, sendo estes apresentados ao longo de uma linha vertical com valor decrescente de similaridade. Assim, as fotografias que apresentam maior semelhança são apresentadas no topo, em tamanho maior, sendo que, à medida que a similaridade diminui, as restantes fotografias vão-se aproximando progressivamente do fundo da linha de similaridade. Esta técnica de visualização, que dispensa a comparação de valores numéricos de similaridade, complementa a interface ao fornecer uma perspectiva visual desta medida.

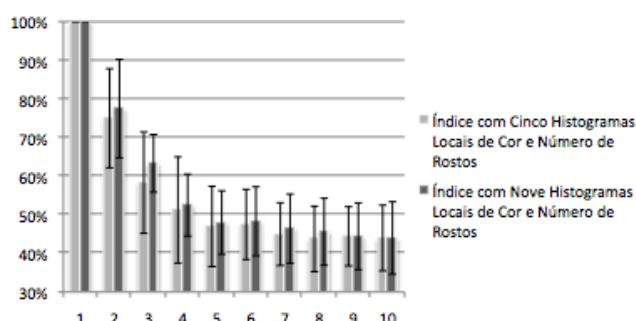
#### 4. TESTES

Neste momentos realizamos apenas testes para comparar as várias abordagens relativas ao algoritmo de recuperação de fotografias. Para isso utilizamos uma base de dados com 1000 fotografias e a R-Precision como medida de desempenho. Esta medida calcula a precisão para cada resultado considerado relevante.

Seleccionamos de forma aleatória 20 imagens da base de dados para usar como interrogações e depois submetermos aos dois algoritmos em avaliação e calculamos a R-Precision média para cada um dos 10 valores mais relevantes (Figura 4).

Da Figura 4 podemos ver que o algoritmo baseado na regra dos terços apresenta resultados semelhantes aos do algoritmo tradicional, com a vantagem de utilizar menos histogramas (cinco contra nove). Isto reflecte-se quer no espaço de armazenamento necessário quer no tempo de processamento das interrogações.

Pretendemos, num futuro próximo, comparar com o histograma global e ainda utilizar uma base de dados de fotografias onde a regra dos terços seja respeitada em todas as fotografias para verificar se os resultados são afectados.



**Figura 4. R-Precision média para os dois algoritmos de recuperação**

#### 5. CONCLUSÕES E TRABALHO FUTURO

A solução que apresentamos consiste numa ferramenta para pesquisa interactiva de fotografias que utiliza histogramas multifacetados e linhas de similaridade como técnicas de visualização. Fizemos ainda um estudo comparativo da utilização de histogramas locais para recuperação de fotografias e obtivemos indícios de que a regra dos terços será uma mais valia na recuperação de imagem. Depois de validados os algoritmos de recuperação, pretendemos efectuar testes com utilizadores para promover um desenho interactivo e a validação da interface.

#### Referências

- [Belongie 01] S. Belongie, C. Carson, H. Greenspan, e J. Malik. Color-and texture-based image segmentation using em and its application to content-based image retrieval. Em *Int. Conf. on Computer Vision (JCCV'01)*, 2001.
- [Fonseca 03] Manuel J. Fonseca e Joaquim A. Jorge. Indexing high-dimensional data for content-based retrieval in large databases. Em *Int. Conf. on database systems for advanced applications (DASFAA' 03)*, 2003.
- [Lee 09] B. Lee, G. Smith, G. Robertson, M. Czerwinski, e D. S. Tan. Facetlens: exposing trends and relationships to support sensemaking within faceted datasets. Em *Int. Conf. on Human factors in computing systems (CHI '09)*, 2009.
- [Rul10] Rule of thirds. <http://digital-photography-school.com/rule-of-thirds>, Accessed June 2010.
- [Smith 97] J. R. Smith e S.-F. Chang. *Querying by color regions using the VisualSEEK content-based visual query system*. AAAI Press, 1997.
- [Stojanovic 07] I. Stojanovic, S. Bogdanova, e M. Bogdanov. Content-based image retrieving improved by pixel-based search. Em *International Conf. on Systems Signals and Image Processing IWSSIP*, 2007.
- [Viola 01] P. Viola e M. Jones. Rapid object detection using boosted cascade of simple features. Em *Conf. on Computer Vision and Pattern Recognition (CVPR '01)*, 2001.
- [Wyszecki 82] G. Wyszecki e W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Wiley, 1982.

# Exploração Interactiva de Colecções Musicais usando Treemaps Ordenados Semanticamente

Ricardo Dias Manuel J. Fonseca  
 Departamento de Engenharia Informática  
 INESC-ID/IST/Universidade Técnica de Lisboa  
 R. Alves Redol, 9, 1000-029 Lisboa, Portugal  
 ricardo.dias@ist.utl.pt, mjf@inesc-id.pt

## Resumo

Neste artigo apresentamos o MuVis, uma nova abordagem para visualização e exploração interactiva de colecções de música, com base no conteúdo musical e informação semântica. A solução desenvolvida é constituída por três componentes principais: componente de visualização de informação (baseado em treemaps ordenados semanticamente), mecanismos de recuperação de informação musical (para a extracção de informação semântica e do conteúdo musical), e componente de interrogações dinâmicas, que oferecem aos utilizadores um modo mais eficiente, flexível e fácil de usar para navegar colecções de música e criar playlists. Os resultados preliminares revelaram que a nossa solução é mais rápida e fácil de usar que o Windows Media Player, e permite aos utilizadores realizarem uma exploração simples e eficiente, enquanto obtêm simultaneamente um maior conhecimento sobre as suas colecções.

## Palavras-Chave

Exploração de colecções musicais, Visualização, Filtragem Interactiva, Criação automática de Playlists

## 1 INTRODUÇÃO

Com o crescimento da popularidade de música em formato digital durante a última década, também a necessidade de mecanismos eficientes e fáceis de usar para a exploração e visualização destas colecções tem aumentado. Embora inúmeras soluções tenham sido desenvolvidas para explorar e organizar as colecções, estas apresentam algumas limitações quando lidam com colecções grandes, tornando tarefas típicas como pesquisar e navegar difíceis. Além disso, também a visualização da informação extraída das músicas têm mostrado pouco progresso e inovação nos últimos anos. A flexibilidade oferecida pelos mecanismos de criação de *playlists* também precisa de ser melhorada, através do uso de informação semântica mas também do próprio conteúdo das músicas.

Alguns dos trabalhos mais recentes exploram técnicas de visualização como o *Self-Organizing Map (SOM)* [Pampalk 01], grafos (Musicoverly<sup>1</sup>), *treemaps* [Torrens 04] e outras abordagens mais inovadoras [Lillie 08] para visualizar colecções musicais. No entanto, a nossa solução difere das demais, na medida em que fornece uma abordagem mais ampla para visualização de bibliotecas musicais, que combina uma visão geral da colecção, um componente de filtragem interactiva e um mecanismo de criação de *playlists* que usa informação

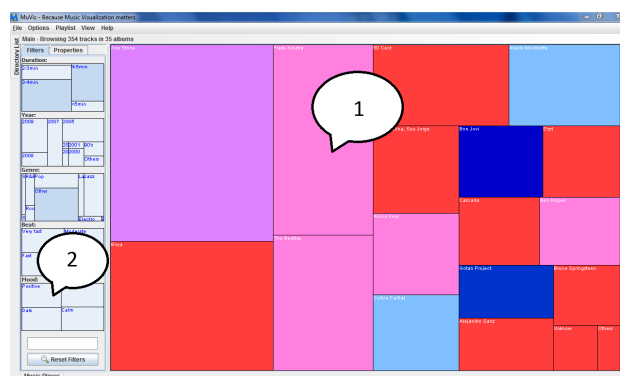


Figura 1. Interface do MuVis.

semântica e do conteúdo musical.

Neste artigo descrevemos uma nova abordagem para a exploração e visualização de grandes colecções de música, que permite aos utilizadores de uma forma fácil e eficiente navegar, explorar e visualizar as suas colecções musicais, bem como criar *playlists*. A nossa solução usa tanto informação semântica como do conteúdo das músicas para fornecer uma visão global da colecção e uma exploração e filtragem interactiva (ver a Figura 1).

Na Secção 2 apresentamos o estudo realizado para recolher informações sobre os comportamentos dos utilizado-

<sup>1</sup><http://musicoverly.com/>, acedido em Setembro de 2010.

res quando ouvem música. Em seguida descrevemos a nossa abordagem, detalhando a interface utilizador e o mecanismo de filtragem interactiva. Na Secção 4 apresentamos os resultados da avaliação preliminar com utilizadores.

## 2 ESTUDO COM UTILIZADORES

Embora alguns estudos tenham sido realizados para compreender as necessidades dos utilizadores enquanto ouvem música [Cunningham 03], estes deixaram de lado o objectivo de entender como os utilizadores executam tarefas típicas nestas aplicações. Assim, antes de desenvolver a nossa solução, começámos por realizar um estudo para compreender os hábitos dos utilizadores quando exploram as suas colecções de música.

Este estudo foi composto por um questionário *online* e por sessões presenciais. Para ambos, criámos um questionário para descobrir como os utilizadores encontram músicas, exploram as suas bibliotecas e criam *playlists*.

Recebemos 127 respostas ao questionário *online*, sendo 65% dos inquiridos do sexo masculino, e a maioria (81%) com idades entre os 20 e 29 anos. Quase todos (94%) ouvem música nos seus computadores pessoais e possuem uma biblioteca de música digital (84%). O tamanho mais típico das colecções situa-se entre as 1.000 e 10.000 faixas (40%), com apenas 17% dos utilizadores com mais de 10.000 faixas, sendo estas no entanto pouco variadas, com mais de dois terços das músicas dos géneros *Pop* e *Rock*. Assim, acreditamos que o género não constitui um bom critério para organizar e estruturar as colecções musicais.

Os resultados não revelaram diferenças entre o uso do mecanismo de pesquisa das aplicações áudio (47%) e o sistema de ficheiros (45%) para procurar músicas, o que nos leva a inferir que as soluções actuais não se distinguem claramente da estrutura hierárquica definida pelo sistema de ficheiros. No entanto, a maioria dos participantes (54%) indicou que gostaria de procurar músicas utilizando *tags*, 48% utilizando o estado de espírito (*mood*) e 46% usando a similaridade entre músicas. Em relação à navegação nas colecções identificámos uma ordem clara pela qual os utilizadores exploram as colecções: Artista (46%), Álbum (41%) e Género (41%). Quanto à criação de *playlists*, identificámos que estas são essencialmente criadas através da selecção individual de faixas (60%), 10% utilizando filtragem automática e *playlists* inteligentes e 11% aleatoriamente. Os participantes indicaram ainda que gostariam de usar o género (51%), *tags* (37%) e o *mood* (37%) como principais critérios para criar as *playlists*.

Para as sessões presenciais, foram seleccionados 10 utilizadores de ambos os sexos com idades compreendidas entre os 20 e os 40 anos, com colecções de cerca de 1.000 músicas e que ouvem frequentemente música no computador. Cada sessão consistiu na execução de um conjunto de tarefas típicas na aplicação preferida dos utilizadores e na resposta ao mesmo questionário. Os resultados desta parte serviram principalmente para validar os obtidos a partir dos questionários *online*.

A partir dos resultados identificámos três implicações de desenho: 1) **Navegação**: A exploração da biblioteca de música deve começar pela selecção do Artista, Álbum, e finalmente Género, 2) **Playlists**: Os utilizadores gostam de deter o controlo sobre a criação das *playlists*; assim, deveremos disponibilizar mecanismos que usam *tags* e o conteúdo musical para os ajudar a seleccionar as músicas desejadas, 3) **Pesquisa**: Para satisfazer as necessidades dos utilizadores, devemos oferecer pelo menos três tipos de pesquisa: i) usando *tags* (por exemplo, *tags ID3*); ii) seleccionando o *mood* (por exemplo, calmo, energético, etc.) e, finalmente, iii) pela similaridade entre músicas.

## 3 SOLUÇÃO DESENVOLVIDA

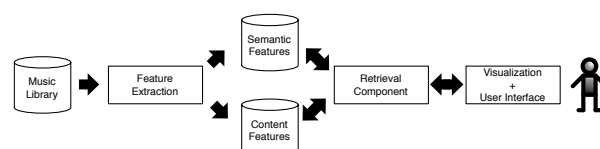
### 3.1 Objectivos de desenho

Os principais objectivos da nossa abordagem são facilitar a exploração de grandes colecções de músicas e a criação de *playlists*. Para atingir estes objectivos, o MuVis oferece uma visão geral sobre a colecção musical, um mecanismo de recuperação baseado em metadados e informação do conteúdo musical, e uma filtragem interactiva para pesquisar e explorar a colecção.

Com a nossa abordagem esperamos ajudar os utilizadores a responder a perguntas como: *De que artista tenho mais álbuns / músicas?*, *Que músicas Rock dos anos 80 tenho na minha colecção?* ou *Quais são as músicas mais semelhantes a esta música?*.

### 3.2 ARQUITECTURA

A arquitectura da nossa solução (ver a figura 2) é composta por um módulo de extracção de informação, um componente de recuperação e a interface utilizador.

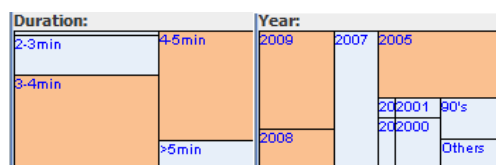


**Figura 2. Arquitectura da solução desenvolvida.**

O módulo de extracção de informação é responsável pela extracção de informação semântica (*tags*, como o nome da faixa, duração, etc.) e informação do conteúdo áudio (*fluctuation patterns* [Pampalk 01] calculados utilizando o CoMIRVA<sup>2</sup>). Para além desta informação, para cada música, um trecho de curta duração (10-15 segundos) representativo da música é extraído.

A recuperação de informação é realizada para os tipos de informação extraídos: *tags*, armazenadas numa base de dados relacional, e informação do conteúdo áudio, armazenada numa estrutura de indexação multidimensional, a NB-Tree. Decidimos utilizar uma estrutura de indexação

<sup>2</sup><http://www.cp.jku.at/comirva/>, acedido em Setembro de 2010.



**Figura 3. Exemplo de dois filtros: duração e ano. Os rectângulos laranja representam os valores seleccionados.**

deste tipo, porque a informação extraída é multidimensional (1200 dimensões) e necessitamos de realizar pesquisas KNN (*K-Nearest Neighbor*) eficientes nesse espaço.

As músicas são apresentadas utilizando duas técnicas de visualização. Uma lista convencional e uma vista de *treemap* ordenada por similaridade musical. Apresentamos em seguida uma descrição detalhada da interface e dos principais elementos.

### 3.3 INTERFACE DA SOLUÇÃO

A interface da nossa solução tem três elementos principais: as duas vistas sobre a colecção (área 1) e o componente de filtragem interactiva (área 2), como ilustrado na Figura 1.

A área 1 corresponde ao elemento principal da interface, o local onde as músicas da colecção são apresentadas aos utilizadores através de duas vistas: vista de lista e de *treemap*. Para a vista de *treemaps* utilizámos uma versão modificada do algoritmo apresentado em [Wood 08]. Nesta vista, os nós da árvore (que representam artistas) são organizados pela sua similaridade face ao artista pivô que corresponde ao elemento localizado no canto superior esquerdo. Os utilizadores podem seleccionar o artista pivô em qualquer momento, e reorganizar o *treemap*. As propriedades dos nós, tais como a dimensão, cor e disposição podem ser controlados através da interface. A dimensão dos rectângulos pode ser a mesma para todos, ou proporcional ao número de faixas / álbuns de cada artista. Os rectângulos podem ser personalizados com cores (que representam o género do artista, de acordo com o código de cores utilizado no sistema *Musicover*), capas dos álbuns ou ambos. Nesta vista os utilizadores podem fazer *zoom*, seleccionando um artista e assim ver os seus álbuns e músicas.

A área 2 contém o mecanismo de filtragem interactiva, descrito na subsecção seguinte.

### 3.4 FILTRAGEM INTERACTIVA

As interfaces para interrogações dinâmicas [Ahlberg 92] permitem aos utilizadores formular interrogações e obter um retorno imediato. Na nossa solução incorporámos esta técnica para oferecer aos utilizadores um mecanismo interactivo de filtragem (ver Figura 1 - área 2), que lhes permite de forma dinâmica filtrar e explorar as suas músicas, mediante a combinação dos vários filtros disponíveis.

O MuVis disponibiliza sete filtros: texto simples, similaridade musical (música, álbum e artista), duração, ano, género, ritmo e estado de espírito. A escolha destes filtros

é consequência directa dos resultados retirados do estudo com utilizadores. O uso de *treemaps* para implementação deste componente, fornece não só uma pista visual sobre a distribuição das músicas por filtro (com base no seu tamanho), mas também a possibilidade de combinar vários filtros interactivamente, permitindo aos utilizadores executar interrogações como as descritas na subsecção 3.1. Cada filtro é representado internamente por um conjunto de regras que definem as restrições a serem aplicadas aos dados. As regras combinam tanto informação da base de dados relacional (*tags*), como da estrutura de indexação (conteúdo), permitindo uma filtragem rápida e mais rica.

### 3.5 CRIAÇÃO AUTOMÁTICA DE PLAYLISTS

A nossa solução permite aos utilizadores criarem *playlists* utilizando as informações extraídas das músicas.

Este mecanismo consiste numa adição contínua de músicas à lista actual, baseada na similaridade com as músicas existentes na lista. Os utilizadores começam por seleccionar um grupo de músicas para servir como sementes e em seguida, o algoritmo adiciona novas faixas sempre que a execução da *playlist* se aproxima do fim. As músicas a adicionar são pesquisadas utilizando interrogações KNN, procurando faixas semelhantes às sementes especificadas. Este mecanismo maximiza o número de músicas semelhantes diferentes que são adicionadas à lista, permitindo aos utilizadores descobrir faixas e manter uma reprodução contínua. No entanto, os utilizadores podem apagar músicas indesejadas da lista, impedindo-as de serem usadas como sementes.

## 4 AVALIAÇÃO

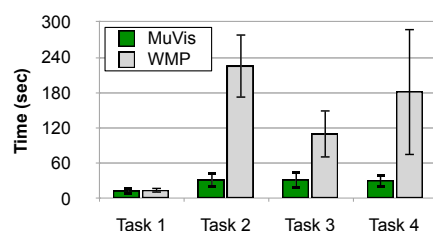
Para validar a nossa solução, realizámos uma avaliação comparativa com o Windows Media Player. Participaram 10 utilizadores de ambos os sexos, com idades entre os 19 e os 40 anos. Todos os participantes possuíam colecções de música e usam o computador diariamente. Os testes foram realizados de forma individual usando o mesmo computador, com o MuVis e o Windows media Player (WMP) carregados com uma colecção de 5.000 músicas de diferentes géneros. Desta forma, colocámos os utilizadores ao mesmo nível, permitindo-nos ter uma base de comparação dos resultados obtidos.

A avaliação experimental consistiu em quatro fases com a duração média de 40 minutos. Em primeiro lugar, descre-

**Tabela 1. Tarefas utilizadas na avaliação.**

Tarefas	Descrições
T1	Encontre a música "Wonderful Tonight" do Eric Clapton e ouça-a.
T2	Crie uma <i>playlist</i> com 12 músicas Pop dos anos 90, com duração entre 3 e 5 minutos.
T3	Diga o nome do artista com mais músicas e com mais álbuns.
T4	Encontre álbuns similares ao primeiro álbum do artista com mais músicas.





**Figura 4. Tempo médio de execução das tarefas para cada aplicação, e o correspondente desvio padrão.**

vemos os objectivos da avaliação e introduzimos as duas aplicações. Em seguida, houve um período de treino, onde os utilizadores usaram livremente as aplicações. Após este período, os utilizadores realizaram um conjunto de tarefas (ver tabela 1) nas duas aplicações. Para evitar o enviesamento dos resultados, metade dos utilizadores começaram os testes pela nossa solução, enquanto os restantes começaram pela utilização do WMP. No fim, os utilizadores responderam a um questionário de satisfação.

#### 4.1 RESULTADOS

No geral, os participantes realizaram as tarefas pedidas mais rapidamente na nossa solução do que no WMP. Além disso, os participantes compreenderam bem o funcionamento do mecanismo de filtragem baseado em *treemaps* e gostaram da pesquisa por similaridade musical.

O tempo médio de execução das tarefas foi menor na nossa aplicação (dois minutos) do que no WMP (oito minutos). Um teste-t realizado revelou que o tempo médio de realização das tarefas utilizando o MuVis (*média* = 109s, *SD* = 37.0) foi significativamente menor do que usando o WMP (*média* = 532s, *dp* = 153.2,  $p < 0.001$ ).

Em média, os participantes realizaram as tarefas 2, 3 e 4 mais rapidamente no nosso sistema que no WMP (ver a figura 4). O teste-t mostra que o tempo médio gasto para cada tarefa é consideravelmente menor no MuVis do que no WMP, com um valor de  $p < 0.002$  para as três tarefas. Para a tarefa 1, embora em média os utilizadores tenham demorado menos tempo no MuVis (13,4 s) do que no WMP (14,5 s) não conseguimos inferir sobre a rapidez da nossa solução, uma vez que  $p < 0.32$ .

Estes resultados mostram que a nossa solução tem um desempenho semelhante ao Windows Media Player para tarefas que envolvem pesquisa textual (T1), mas tem um desempenho superior em tarefas que envolvam navegação e exploração da colecção musical (T2 e T3) e pesquisa por similaridade musical (T4).

A maioria dos participantes (quase 80%) manifestaram grande satisfação com a nossa solução, principalmente porque oferece uma navegação simples da colecção musical, utilizando um mecanismo de filtragem interactivo com uma representação visual, que permite localizar músicas eficientemente. A criação de *playlists* também foi descrita

como mais fácil (60%) devido à sua flexibilidade, uma vez que é possível criar uma *playlist* a partir das várias vistas da aplicação.

Finalmente, observámos que os utilizadores foram capazes de obter uma visão correcta do conteúdo da colecção musical, identificando os principais géneros (Pop, Rock, R&B), os artistas com mais álbuns e com mais músicas. Os participantes mencionaram ainda que o tamanho dos diferentes rectângulos os ajudou a ter uma boa percepção da colecção musical sobre as várias perspectivas que estes representam.

#### 5 CONCLUSÕES

Neste artigo apresentámos uma nova solução para a exploração interactiva de grandes colecções musicais. A abordagem desenvolvida fornece uma forma rápida e simples de interagir com a biblioteca musical. Os resultados experimentais mostraram que o nosso objectivo de desenvolver uma ferramenta eficiente para a visualização e exploração de colecções musicais foi alcançado, satisfazendo os utilizadores.

Futuramente, pretendemos desenvolver e aplicar novas técnicas de visualização, como as técnicas de foco+contexto introduzidas em [Tu 08], e ir mais longe na avaliação da solução proposta, definindo novas tarefas e utilizando mais participantes, de forma a avaliar as consequências da utilização da similaridade na visualização em *treemaps*.

#### Referências

- [Ahlberg 92] Christopher Ahlberg, Christopher Williamson, e Ben Shneiderman. Dynamic queries for information exploration: an implementation and evaluation. Em *CHI*, 1992.
- [Cunningham 03] Sally Jo Cunningham, Nina Reeves, e Matthew Britland. An ethnographic study of music information seeking: implications for the design of a music digital library. Em *JCDL*, 2003.
- [Lillie 08] Anita Shen Lillie. Musicbox: Navigating the space of your music. Tese de Mestrado, MIT, 2008.
- [Pampalk 01] Elias Pampalk. Islands of music: Analysis, organization, and visualization of music archives. Tese de Mestrado, Vienna University of Technology, 2001.
- [Torrens 04] M. Torrens, P. Hertzog, e J. L. Arcos. Visualizing and exploring personal music libraries. Em *ISMIR*, 2004.
- [Tu 08] Ying Tu e Han-Wei Shen. Balloon focus: a seamless multi-focus+context method for *treemaps*. *TVCG*, 2008.
- [Wood 08] Jo Wood e Jason Dykes. Spatially ordered *treemaps*. *TVCG*, 2008.

## **Sessão VI**

### **Videojogos e Interacção**

Aveiro, 13, 14 e 15 de Outubro de 2010



# Measuring Difficulty in Platform Videogames

Fausto José Mourato  
Departamento de Sistemas e Informática  
Escola Superior de Tecnologia  
Instituto Politécnico de Setúbal  
2910-761 Setúbal  
fausto.mourato@estsetubal.ips.pt

Manuel Próspero dos Santos  
CITI, Departamento de Informática  
Faculdade de Ciências e Tecnologia  
Universidade Nova de Lisboa  
2829-516 Caparica  
ps@di.fct.unl.pt

## Abstract

*Automatic generation of game levels improves replayability and also allows content adaptation. One important aspect to take into account in the creation of any videogame is difficulty, in particular when it is possible to adapt content. However, defining difficulty is not a straight subject. In this paper we propose a metric for game difficulty in platform levels, mostly based on users' losing probability for each obstacle. This metric can be further used in automated processes that generate levels for this type of games, helping the process to recognize whether a level is suitable or not for a certain player. We also present some examples of the usage of this metric in commercial games.*

## Keywords

*Human Factors, Platform videogames, difficulty measurement.*

## 1. INTRODUCTION

Procedural generation of content is one active area in Computer Graphics with some common associations to other areas such as Artificial Intelligence and Human-Computer Interaction. In this paper we direct our main focus to the automated procedural generation of levels for platform games, aiming the contributions to understanding the related concept of difficulty. Considering a system that automatically generates platform levels, estimating and quantifying difficulty is useful because it improves the knowledge about the output.

The designation of platform games is used to describe a particular type of videogames where the user controls an avatar that has to accomplish a set of challenges. These typically consist of jumps between elements which are referred as platforms. Some popular examples of this type of games are: *Super Mario Bros*, *Sonic – The Hedgehog* and *Little Big Planet*. Screenshots of these three games are provided in Figure 1.

The main difference in generating automatically platform levels in comparison with other automated graphical generations, such as buildings, trees or even terrain, is that

the final result represents a challenge with a certain difficulty, and a minor change can affect it with indefinite proportions. For instance, one platform moved slightly from its original place in a well balanced level can make it impossible to finish. Therefore, the main question that arises is: How do we measure difficulty in a platform level?

In order to answer the previous question, we propose a method to measure difficulty in a platform level, which we stated that can be considered as a relation to the probability of failure. Inside this topic our main contributions are:

- A study on the main questions related to difficulty in platform games.
- A proposal for a method to measure difficulty in platform game levels, relying on success probability.
- A practical usage of the difficulty proposed principle on some levels of existing commercial applications.

In the next section we will explain some of the major motivations behind difficulty measurements in the topic of automated generation of videogame content.

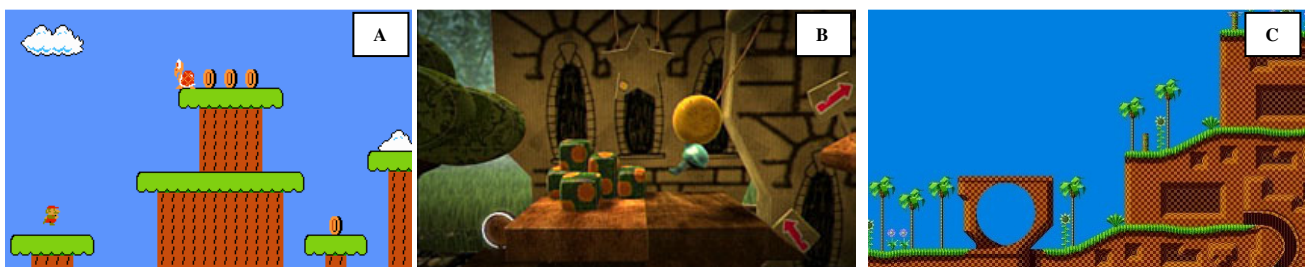


Figure 1 – Screenshots of the games *Super Mario Bros.* (A), *Little Big Planet* (B) and *Sonic – The hedgehog* (C)

## 2. WHY MEASURE DIFFICULTY?

As stated in the previous section, this study consists in improving the knowledge about difficulty in platform games, which can represent a step forward in processes that automatically generate this kind of levels. There are several possible approaches for platform level generation, but only a few have been made effective, as we will see in a later stage of this document, in section 5, where the related work is presented with more detail.

To better understand the problem, we can consider a system that generates levels on-the-fly where the user can play. For the sake of ease we can also consider by now that the system validates physical restrictions and thus produces always valid levels. This means that the user will always have a chance to succeed, depending only on his/her skills. However, even in these conditions, there is no way to ensure that the challenge is appropriate for the user because all levels are treated equally. This leads us to the need of matching users' skills to provided challenges, which should have an appropriate dimension of difficulty. As a result, it is required that generated levels have a certain difficulty value. In the next section we will look at the most common situations that are present in most platform games in order to understand where the difficulty lies and hence measuring it. In section 4 we will also present some examples to show that several common game situations can be fitted in our reference situations.

## 3. DIFFICULTY IN PLATFORM GAMES

### 3.1 Understanding difficulty in a whole level

We presented the problem of measuring difficulty on platform game levels and its importance for automated procedural generation, in particular for content adaptation. In this section we will try to answer the following question: What makes a level difficult?

In order to understand where difficulty lies, our main examples will consider the more trivial cases for platform levels and their possible adaptations, covering the common structures that this type of games have, as proposed by Compton and Mateas [Compton06]. Levels are defined with one start point and one end point in a bi-dimensional geometry. Some levels may have certain objects that harm the player's avatar causing him to lose. These will be represented as spikes, which are common objects in this type of games. In addition, there are two main considerations to take into account regarding to the user experience, which are the following:

- The users are aware of the task they are doing, which means that every challenge is directly identifiable, as proposed by typical usability heuristics such as those presented by Desurvire et al. [Desurvire04].
- Trivial tasks, such as moving from one place to another in a single platform or climbing up a ladder, are not relevant in difficulty measurements.

With those principles we will consider the five different cases represented in Figure 2, which can be briefly described as follows:

- A straight way where it is impossible to fail. Since the user knows what to do he/she will not run blindly against the left wall.
- An impossible level because the end point is physically unreachable.
- An almost impossible level with the end point being almost physically inaccessible.
- One level with a single jump that the user must accomplish to reach the end point. Failure attempting the jump will cause the player to lose.
- A level where the user needs to jump from one block to another but if he/she fails he/she will be able to try again after passing through another obstacle that may cause him to lose. The stairs represented in this case simply mean that climbing to the first platform is a trivial action that every user can do without difficulties.

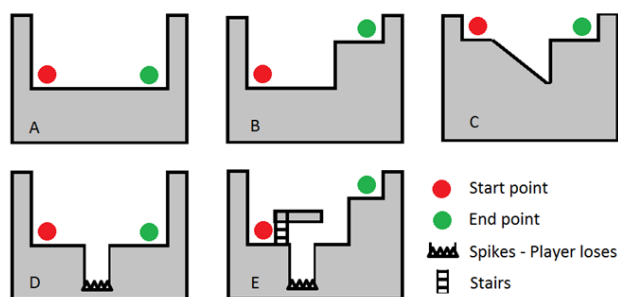


Figure 2 – Five examples of common situations in platform games

The first case is a generalization to represent the easiest level that can exist. Due to the subjective nature of the topic this cannot be proven but common sense seems to be enough for this situation. The question here can be reversed and be thought with the following question: Can we design an easier level than this? In fact, the task required to complete the referred challenge can be described as a single action, where there is no possible way of losing. Simplifying the problem, it is plausible to say that this first example is easy because the probability of losing is zero. Completing this level consists on doing one simple task with a probability value that can be considered one. A simple graphical representation is available in Figure 3.

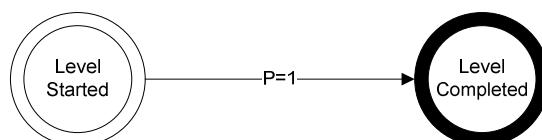
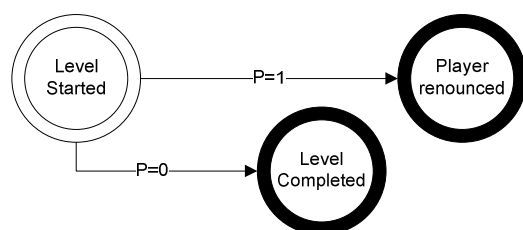


Figure 3 – State chart for example level 1

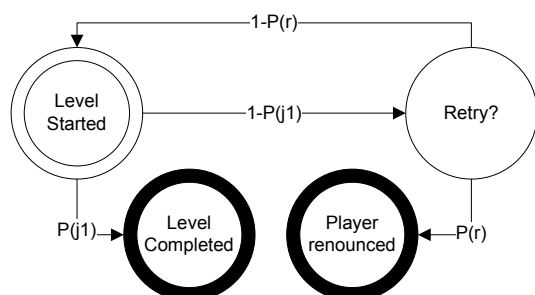
The second example tries to create a contradiction with the conclusion that we can retrieve from the first example. In this case, the probability of losing is again zero, because there are no elements to harm the player's avatar. However, winning probability is also zero because the user needs to jump to an unreachable place. The same problem will occur if we think of an adaptation of the first example in which the start and end positions are too

far apart that there is no user with enough patience to complete the challenge. This leads us to another aspect to take into account that is the probability of renouncing. Considering the psychological approaches to this subject, the user will resign playing if he/she is bored or frustrated. Boredom will occur in a situation as the adaptation of the first example, with an almost infinite distance between start and end points. Frustration will occur in situations as the second example, where the user is presented with an impossible task. The idea of a good game is to keep the user in-between those two states, in a state that is designated as flow [Csikszentmihaly91]. The state chart in Figure 4 is a possible representation of the described situation.



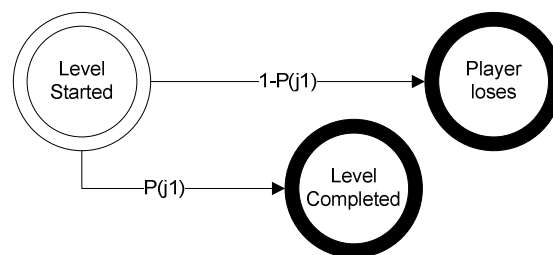
**Figure 4 – State chart for example level 2**

In the third example we keep the user again in a harmless level with one single jump that has to be accomplished. The main difference is that in this case the full probability is divided in success and giving up. A state chart is provided in Figure 5 to represent this situation, considering  $P(j_1)$  the probability of success in the first obstacle, in this case a jump, and  $P(r)$  a general resign probability for the user after failing one challenge. If we consider an impossible jump, it is possible to state that this case and the correspondent diagram will result in the previous situation. In the next subsection we will explore the associated calculations that can be extracted.



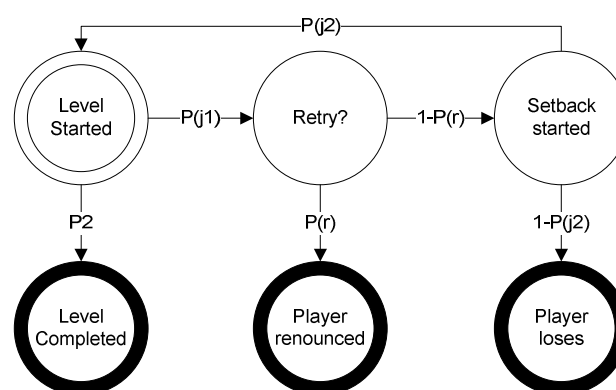
**Figure 5 – State chart for example level 3**

The fourth example adds the existence of harm, which cause the player to fail immediately. Considering that the distances between the start and the end point are small, we can consider that the user will not forfeit because the level consists in one single task. So, the probabilities will be divided directly in winning and losing, as represented in Figure 6. Those probabilities are only linked with the difficulty of the jump, which will have a probability of success defined by  $P(j_1)$  as in the previous example. We will also approach this problem with some more detail in the next subsection.



**Figure 6 - State chart for example level 4**

Finally, the fifth example gives us a single jump with no direct harm that causes failure but that requires succeeding in another jump in order to retry. The result is a loop as in the previous case. The main difference is that, in this last example, the retrying loop can be broken either by renounce or by failing in the process. This is represented with the state chart in Figure 7.



**Figure 7 – State chart for example level 5**

After the previous examples we can identify two main variables that contribute to difficulty in one level:

- $P(r)$ : The probability of renouncing after failing one attempt, which we will consider constant in one level. This value is directly related to the player's resilience.
- $P(j_i)$ : The probability of achieving success in the jump or similar challenge with the index  $i$ .

Also, we could identify that there are two main elements to consider for the difficulty calculations over one level, which are the obstacles where the avatar may directly lose and the obstacles that will make the avatar repeat other obstacles that may eventually cause renounce or failure.

With these concepts in mind, we will proceed to the next subsection to understand how to extract final probability values and thus difficulty.

### 3.2 Measuring difficulty in a level

In the previous subsection we represented the most common situations in platform games and identified that difficulty is related to the probability of being unsuccessfully, which might be caused by failure or resignation. The studied situations were represented as state charts in which every transition is associated with a probability value. Our proposal is to calculate the probability of suc-

cess and failure that is recursively represented in those charts, and thus extract difficulty.

Calculating probabilities in cases like one, two and four are simply a sequence of independent cases of probabilities, thus measuring difficulty can be calculated by multiplying all the values from the start to the end point. Cases like the examples three and five have loops in the representation, which makes the calculation less simple. For instance, probability of success in example four is the probability of making the jump on the first attempt plus the probability of making it at the second attempt, if there is a second attempt, and so on. This can be expressed mathematically with an infinite equation to represent the probability of success  $P(s)$  as follows:

$$P(s) = P(j1) + P(j1') \cdot P(r') \cdot P(j1) + \dots + (P(j1') \cdot P(r'))^n \cdot P(j1) + \dots$$

This equation can be compressed, resulting in the following:

$$P(s) = P(j1) \sum_i (P(j1') \cdot P(r'))^i$$

If we think of a whole level, measuring these probabilities with a formal model will require a long set of equations and the calculation of these series. However, the information is more likely to be represented as a state-chart or a graph rather than in mathematical notation, so these calculations might be done with a recursive and iterative algorithm that spreads probabilities over the nodes until a certain precision is reached.

Considering the previous representations for the whole problem and that we can recursively spread probability values through a graph in order to understand what are the final probabilities to achieve success and failure in one level, it is important to extract the probability values for the connections in the graphs, which represent the difficulty for isolated jumps or, more generically, isolated obstacles. This topic is about to be approached in the next subsection.

### 3.3 Difficulty for independent obstacles

In this subsection we propose a method for measuring the difficulty for each isolated jump or challenge, which can be used to feed the probability graphs presented in the last subsection.

One simple assumption is that a single jump is more difficult to accomplish if the gap to jump is bigger. This is valid for platforms at the same height, but for different heights the problem is not that direct. With a physics simulator it is possible to identify the multiple alternatives to make a successful jump between two platforms. Compton and Mateas [Compton06] already proposed this approach but only as a theoretical possibility without concretization. In addition, looking at all existing combinations of jumps also gives just a notion on the difficulty of a jump by counting the successful and failure cases. Some of those cases will be more likely to happen than others, because the user keeps correcting the movement in the jump. Since transposing physical constraints to a difficulty measurer may be a difficult task, especially if we are estimating difficulty in an existing game with

closed source, we propose a faster alternative based on the jump characteristics. So, in a jump, at a higher level of abstraction, the player is trying to jump from one point to another. These points represent the edges of the origin and destiny platform. With this in mind, our approach consists in launching a projectile from the first point  $P_0$  and measuring a possible margin of error for its trajectory relatively to the second point  $P_1$ .

Recalling some basic physics, a projectile trajectory can be defined by the following set of equations in order of time:  $x(t) = x_0 + v_{0x} \cdot t$  and  $y(t) = y_0 + v_{0y} \cdot t - \frac{1}{2} a \cdot t^2$

Once every jump is relative to the platform from which the user is jumping, we can consider the origin point ( $P_0$ ) as the reference to simplify the equations. Also, we will not consider a throwing angle, so the initial speed can be defined as a constant ( $K_i$ ), configurable for different games. Finally, these calculations are space oriented and the jump duration is not used, so the calculation can be resumed as one expression, which is basically a quadratic equation intersecting the origin, simply defined by:

$$f(x) = K_i \cdot x - \frac{1}{2} a \cdot x^2$$

This trajectory can be seen as the biggest possible jump the avatar can perform. By intersecting the projectile with two lines that are parallel to the Cartesian axis and that also intersect the destiny point, we can estimate the possible deviation to the trajectory that is still a valid jump. From now, this deviation will be referred as the *error margin*. In fact, we are identifying two values: the height of the player when he/she horizontally reaches the platform ( $\Delta y$ ) and the horizontal amplitude of the jump ( $\Delta x$ ). So, our error margin has values in x and y axis ( $m_x$  and  $m_y$ ) defined by  $\Delta x - x_1$  and  $\Delta y - y_1$ , respectively, with  $P_1 = (x_1, y_1)$ . This concept is graphically summarized in Figure 8.

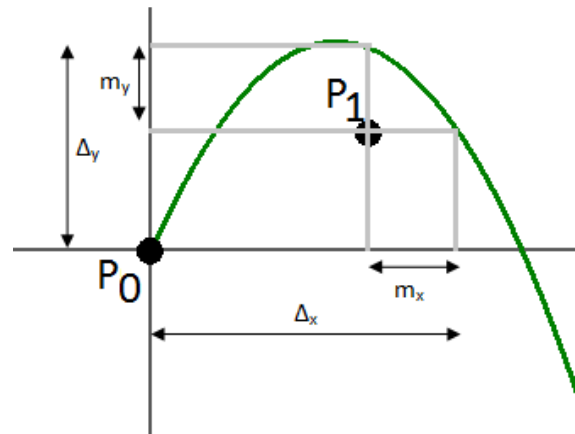


Figure 8 – Graphical representation of the concept of error margin

Obtaining  $m_y$  requires knowing  $\Delta y$ , which consists on calculating  $f(x_1)$ . A negative value means that the platform is unreachable with the reference jump. Here, the system has to identify whether the platform is really unreachable and thus the error margin is immediately zero without any other calculations, or the jump has to be attempted differently (for instance, moving the origin point



to the left). For positive values we measure the horizontal tolerance by calculating the value of  $x$  for which  $f(x)$  equals  $y_1$ , which consists in solving a second degree equation and selecting the appropriate root. To normalize results we consider the error margins as relative values to the full distances measured along the axis. In the end we multiply both the normalized error margins to define the final error margin for the obstacle ( $M$ ), defined between zero and one.

Empirically, one can perceive that difficulty is related to the error margin in a non-linear fashion. As error margin decreases it is expectable to have a few failures occasionally occurring, but as the error margin gets lower, more and more failures are willing to occur. To reflect this principle, we represent the final difficulty value for each obstacle based in an exponential function with the following equation for difficulty in the obstacle of index  $i$ :

$$D(i) = M_i^{Kd}$$

The constant  $Kd$  represents the mapping of the linear error margin and the expected exponential distribution that can be configured to a certain player profile. In the results presented in section 4 we will present some comparisons.

### 3.4 Dynamic entities

Our main principles consider a static scenario, which is somehow a limitation. Some alternatives and proposals for the dynamic question will be addressed in future work. However, in this subsection we will present how we deal with this problem in our current studies.

Typically, in a platform game, dynamic objects are moving platforms and enemy creatures. For the first ones, we can consider a gap that separates two platforms and that is impossible to jump across the gap. One moveable platform slides from one side of the gap to the other, so the user needs to use it as a transport. Considering that the user is focused in the task and it is not exploring the scenario in a way that is not intended, this process consists in jumping two gaps. The main issue is that the movement of the dynamic platform produces an extra error probability. Also, the faster the platform moves, the bigger the error might be. So, for situations like this, we apply the calculations previously presented for normal jumps. However, to represent the increase of difficulty caused by dynamics, the gap distance to consider is a normal jump to the platform at the nearest place plus an extra term that is proportional to platform speed and a constant. Representing enemies can also be considered as jumping implicit gaps. Depending on the enemy's characteristics (mainly speed and size), we manually define a value for  $P_1$ .

### 3.5 Final remarks

In this section we have presented an approach to measure difficulty in platform game levels. For each obstacle we defined an error margin to extract an estimated probability of succeeding. The whole level can be considered as a graph with multiple transitions where the referred probabilities can be applied to obtain a global value. The next section presents some results about applying this concept.

## 4. CASE STUDY

In this section we will briefly explain our experiences on measuring difficulty in levels of existing commercial applications. Our experiments consisted on the following:

- Mapping existing game levels in a representation with only the situations previously presented and measuring difficulty in the mapped levels with the principles previously presented.
- Testing levels with real players on the referred levels, measuring their probability of success and comparing to the estimated results.

### 4.1 Mapping and measuring existing levels

As we previously referred, we applied our proposed metrics to some existing games. The classics are useful because normally it is possible to find online bitmaps with whole level representations. The results of this subsection present difficulty measurement in some levels of the game *Super Mario Bros*.

First, in Table 1 we present the success probability for a regular user in some levels of this game measured as we proposed in this document. We selected a set of levels that share similar objects and obstacles to avoid having particular aspects biasing the conclusions, such as the existence of *bosses*, portals or different physical conditions (for instance, some levels are played under water). The value  $T$  in the table represents the number of times the user will retry an unsuccessful jump.

Level	T=1	T=2	T=3	T=∞
W1L1	10.8%	32.4%	40.8%	44.6%
W1L2	3.5%	10.7%	13.7%	15.0 %
W1L3	.5%	.8%	.9%	1%
W2 L1	.22%	3.15%	5.8%	7.4%
W3 L1	2.4%	15.9%	21.6%	24.0%
W3 L2	3.6%	13.2%	15.3%	15.7%
W3 L3	.40%	.51%	.52%	.52%

**Table 1: Measured values in Super Mario Bros.**

One interesting aspect to notice in these results is that, in fact, levels tend to have higher difficulty values as the game evolves (less probability of success). It is also possible to notice that different resilience values for the user ( $T$ ) influences the final probability of success. In particular, a value of 1 represents the less resilient player that can exist, which is one that resigns automatically after failing one simple jump.

This particular game, as some others of the genre, organises the levels in groups named as worlds (in table 1, W stands for world and L for level). One can also notice that when a new world starts difficulty tends to have a small decrease before rising again. This might represent an intention of having a resting level after completing a difficult task, in this case, the last level of the previous world.

Moreover, tuning the exponential coefficient allows adjusting the values in proportions to have them to be more

reliable for usage as real success probabilities. Lower values are more suitable to represent the more skilled players, which are very unlikely to lose in small challenges. Contrary to this, higher values represent less skilled players that have higher failure rates immediately as challenges appear.

In Table 2 we fixed the value of  $T$  to 3 and vary the exponential coefficient. So, each column represents the probabilities for players with different skills in the levels presented in each row.

Level	Kd=.25	Kd=.33	Kd=.50
W1L1	40.8%	28.9%	12.6%
W1L2	13.7%	6.8%	1.4%
W1L3	.9%	.2%	.01%
W2 L1	5.8%	2.0%	.2%
W3 L1	21.6%	13.3%	3.3%
W3 L2	15.3%	8.2%	2.1%
W3 L3	.52%	.10%	.003%

**Table 2: Measured values for difficulty in Super Mario Bros. with different Kd Values**

We proceed now to some tests we ran with a set of players in the game *Little Big Planet*. This game was particularly useful because it allows users to create their own levels. With this, it was possible to create our own test set. We created a small template level with a few jumps and adapted it to have different versions with distinct gap sizes in order to influence the difficulty. In table 2 we present the predicted and the effective probability of success for each obstacle.

Real P(s)	Pred. P(s) Kd=.3	Pred. P(s) Kd=.2	Pred. P(s) Kd=.1
40,4%	40%	54%	74%
89,3%	48%	62%	78%
91,7%	65%	75%	87%
96,2%	57%	69%	83%
96,2%	57%	69%	83%
96,2%	94%	96%	98%
97,1%	78%	84%	92%
98,0%	74%	82%	90%
98,0%	94%	96%	98%
98,0%	87%	91%	95%
98,0%	89%	92%	96%
98,0%	81%	87%	93%
100,0%	95%	96%	98%

**Table 3: Examples of measured difficulty in the levels created for the game *Little Big Planet***

Even though we have not yet extracted correlations to the obtained values, some relations are detectable. In further

experiments we intend to apply some categorizations to our users, mapped to different values for  $Kd$ , in order to achieve for accurate values.

## 5. RELATED WORK

### 5.1 Platform videogames

Platform games have been studied before by a few authors. There are three main relevant works to refer in the context of this article. In fact, those three articles are the result of continuous work inside the same research group.

In the first work, Compton and Mateas [Compton06] studied the structure of platform levels, identifying some construction parameters to configure platforms. Also, they identified that platforms are associated together by patterns that represent the actions that the avatar has to do in order to pass through a certain section in one level, designated as cell. They also identified the main structures of organising cells in one level, which was particularly helpful in the definition of the examples presented in the previous sections of this document. Finally, they stated that a system that automatically generates levels could be improved by measuring difficulty. As a theoretical approach, the authors pointed this analysis as one step of the process. In their proposal, one possible approach for this was to calculate all possible trajectories for each jump from one platform to another. With this, they wanted to be able to calculate the spatial window to a possible successful jump. In some aspects, the concept of error margin previously presented in this article represents a window like this. Also, they wanted to use the referred physical calculation to extract the time window the player has to make corrections to the movement.

Later, Smith et al. [Smith08] defined a framework to analyse platform levels in which concerns to its structure. Some important concepts were formalised, resulting in a conceptual model to define a generic level. The need of classifying difficulty was identified once again but only proposed as a future work. Recalling the simplifications or abstractions previously considered to deal with dynamic objects and enemies, the same approach was considered by these authors. For instance, they defined a platform level as a hierarchy where everything that may cause damage to the player is a generic obstacle. This means that one gap between two platforms is an obstacle. Since in our work we directed our measurements to distance, our abstraction was built in the other way around, by mapping objects into gaps.

The previously presented concepts lead to an effective implementation of a level generator [Smith09]. This implementation starts with a rhythm generator that creates a set of actions to be done by the avatar. Those actions are then used with a physics system in order to generate a valid geometry, which is one that allows the user to replicate the generated actions. A system based on critics analyses the generated levels to avoid over generation and also to establish a quality threshold. However, the geometry that is created is still not analysed with a difficulty perspective. Once again, it was identified the need of measuring difficulty for further implementations. Also, the authors stated that measuring difficulty can be used

inside the critics system. The approach we have presented allows analysing a level with the condition of having the set of actions well identified. Once their proposal is rhythm-based, which means that all the level is generated based on a set of actions, difficulty could be measured as we proposed.

Besides the three previous works, there is one particular study for platform games that is important to refer, developed by Pedersen et al. [Pedersen09]. Their focus was directed to more abstract concepts about user experience. In particular, the authors studied the users' perceptions in an adapted version of the game *Super Mario Bros.* Several users played the game with different features to establish correlations between those features and reported emotions. A model was built with the obtained results in order to predict fun, challenge and frustration.

## 5.2 Difficulty measurements in other genres

In other genres, difficulty has been used in some more concretized manners. Just to give a brief overview of that, we will point a few examples.

Togelius et al. [Togelius07] presented a system to automatically generate racing tracks for a driving simulator. To evaluate each track quality for usage as fitness function in a genetic algorithm, they used artificial drivers, mapped to a certain profile, to extract some attributes such as timings and speeds.

A similar approach was used for *Pac-Man-like* games by Togelius and Schmidhuber [Togelius08]. The interesting aspect of this work is that the generation process produces game variants instead of game levels. Again, the main principle used was evolutionary computation and the measurements of difficulty were accomplished by making intelligent agents play the game.

Finally, another interesting work to refer was developed by Pereira et al. [Pereira09]. In this case the objectives and approaches are considerably different from the last ones. The considered genre was Strategic Multiplayer Browser Game, which consists on a slow paced evolution system, accessed by players a few times on one day to establish some strategies about virtual resource management. The system tries to involve the player in an ambient that fits the user preferences, with a balanced distribution of resources to avoid repetition. Also, the authors refer the importance in handling efficiently the constant appearance of new players in the earlier stages of the game.

Naturally, there are several other examples of dynamic adjustment of difficulty, in particular on commercial games. Nevertheless, the presented cases are particularly interesting because they have emphasis on automated generation of content.

## 5.3 Practical usage of difficulty measures

As presented in this document, the main importance of quantifying difficulty is that it also allows adapting it. Adapting difficulty in videogames has the main objective of levelling a challenge among intervenients. However, it is important to make those adaptations without compromising the core player experience [Hunnicke05]. For a

single player game, adapting difficulty normally consists on tuning artificial intelligence behind virtual agents. In multiplayer games the key idea is to make the challenge higher to expert players and lower to casual player. The final goal is that every player has a similar chance of winning. Basically, it is the same intent as in some sports that use handicap, such as golf. However, that system is not always applicable, and in some videogames we can make performance analysis in real-time and with that adjust difficulty in a more transparent way. One example of implementation of this concept is the work proposed by Martínez and Mata [Martínez09] in which a *Pong* version is adapted according to players' skills. The authors considered that, in this game, difficulty is due to the paddle size and ball speed. This two attributes are adapted during game play to balance the result and level the odds of each player. In addition, these values are also adapted to make the game challenging enough for both players, and not too easy or hard for both.

## 6. CONCLUSIONS AND FUTURE WORK

### 6.1 Conclusions

We have proposed a way to measure difficulty in platform games and presented its possible application in real games. With basic probability principles, we define the flow of action in a level in association with failing probabilities to obtain a final probability of completing a level with success. Resilience parameters were also considered to reflect the possibility of a user to resign on harder challenges. In addition, basic projectile principles allowed a simple estimation of difficulty for each individual obstacle, which is mapped in a probability value to be used in the calculations of the full success probability.

Our primal experiences show that estimating difficulty based on probability gives an effective notion about its growth. We have tested our measurement in some existing games, in particular a few classics in the gaming industry. As presented previously in section 4, with the results obtained in *Super Mario Bros.*, we could verify that our approach generally identifies the earlier levels as being easier than the later ones. This goes in line with the empirical notion of common users and basic design rules.

One important aspect is that the obtained results don't consider the existence of lives (attempts) and checkpoints. So, the probability values presented estimate the chance that the user has to complete the whole levels in one single try.

Finally, our tests with users also allowed identifying a relationship between the proposed difficulty estimator and the real probability of success. Once again, it is noticeable that the difficulty measured can estimate if a level is easier or harder, but it does not represent an effective probability of success in one level.

### 6.2 Future Work

Even though we have identified some of the principles where difficulty in platform games is based on, some of those need to be analysed with more depth.

Firstly, we believe that it is important to expand the concept of obstacle. The simplifications used for enemies



and dynamics were a sufficient solution for this first approach, but this is a question that requires other calculations to be defined with more detail.

As stated before, the measurements didn't consider the existence of checkpoints and lives that allow the user to retry the level. It would be interesting to expand the study taking this aspect in account and identifying possible differences.

Also, it would be interesting to consider the existence in some levels of the so called *bosses*, which are particularly skilled enemies. For that reason, some levels were not analysed.

Moreover, it is important to make parameter adjustment based in real measurements rather than the ad-hoc tuning that was used. In particular, the coefficient used can represent the user skills, hence predicting more accurately the user's probability of success.

Finally, an interesting way to improve data gathering and analysis, in particular on what concerns to difficulty, might consist in the creation of a community where players could freely play automatically generated levels. This could improve the results obtained in-game.

## 7. REFERENCES

- [Compton06] Compton, K. and Mateas, M. 2006. Procedural Level Design for Platform Games. *Proceedings of the 2nd Artificial Intelligence and Interactive Digital Entertainment Conference* (Stanford, CA, 2006).
- [Csikszentmihaly91] Csikszentmihaly, M. Flow: The Psychology of Optimal Experience. Harper Collins, NY, 1991.
- [Desurvire04] Desurvire, H., Caplan, M., and Toth, J. A. 2004. Using heuristics to evaluate the playability of games. *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria, April 24 - 29, 2004). CHI '04. ACM, New York, NY, 1509-1512.  
<<http://doi.acm.org/10.1145/985921.986102>>
- [Hunicke05] Hunicke, R. 2005. The case for dynamic difficulty adjustment in games. *Proceedings of the 2005 ACM SIGCHI international Conference on Advances in Computer Entertainment Technology* (Valencia, Spain, June 15 - 17, 2005). ACE '05, vol. 265. ACM, New York, NY, 429-433.  
<<http://doi.acm.org/10.1145/1178477.1178573>>
- [Martínez09] Ibáñez-Martínez, J. and Delgado-Mata, C. From competitive to social two-player videogames. *Proceedings of the 2nd Workshop on Child, Computer and interaction* (Cambridge, Massachusetts, November 05 - 05, 2009). WOCCE '09. ACM, New York, NY, 1-5.  
<<http://doi.acm.org/10.1145/1640377.1640395>>
- [Pedersen09] Pedersen, C., Togelius, J., and Yannakakis, G. N. 2009. Modeling player experience in super mario bros. In *Proceedings of the 5th international Conference on Computational intelligence and Games* (Milano, Italy, September 07 - 10, 2009). IEEE Press, Piscataway, NJ, 132-139.
- [Pereira09] Pereira, G., Santos, P. A., and Prada, R. 2009. Self-adapting dynamically generated maps for turn-based strategic multiplayer browser games. *Proceedings of the international Conference on Advances in Computer Entertainment Technology* (Athens, Greece, October 29 - 31, 2009). ACE '09, vol. 422. ACM, New York, NY, 353-356.  
<<http://doi.acm.org/10.1145/1690388.1690457>>
- [Smith08] Smith, G., Cha, M., and Whitehead, J. 2008. A framework for analysis of 2D platformer levels. *Proceedings of the 2008 ACM SIGGRAPH Symposium on Video Games* (Los Angeles, California, August 09 - 10, 2008). Sandbox '08. ACM, New York, NY, 75-80.  
<<http://doi.acm.org/10.1145/1401843.1401858>>
- [Smith09] Smith, G., Treanor, M., Whitehead, J., and Mateas, M. 2009. Rhythm-based level generation for 2D platformers. *Proceedings of the 4th international Conference on Foundations of Digital Games* (Orlando, Florida, April 26 - 30, 2009). FDG '09. ACM, New York, NY, 175-182.  
<<http://doi.acm.org/10.1145/1536513.1536548>>
- [Togelius07] Togelius, J. Nardi, R and Lucas, S. Towards automatic personalised content creation for racing games, *Proceedings of the IEEE Symposium on Computational Intelligence and Games*, 2007.  
<<http://cogprints.org/5573/>>
- [Togelius08] Togelius, J. and Schmidhuber, J. An experiment in automatic game design. *Proceedings of the IEEE Symposium on Computational Intelligence and Games*, 2008.  
<<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.141.4110>>

# Wiisplay: Avaliação da interacção

André F. S. Barbosa      Frutuoso G. M. Silva  
 Instituto de Telecomunicações  
 Universidade da Beira Interior  
 Rua Marquês D'Ávila e Bolama, Covilhã, Portugal  
 andrefsbarbosa@gmail.com      fsilva@di.ubi.pt

## Sumário

*Este artigo descreve um videojogo 3D de plataformas que permite jogo individual ou cooperativo e que utiliza detecção de movimentos como forma de interacção com o jogador. O jogo foi construído através da plataforma XNA e utiliza o comando da Nintendo Wii, o Wiimote. Neste artigo são discutidas vantagens da utilização de detecção de movimentos em videojogos, assim como a sua aplicação em Serious Games. São apresentados também alguns resultados do inquérito feito para avaliar a utilização de movimentos para controlar o jogo, em comparação com os periféricos tradicionais.*

## Palavras-chave

*Serious Games, Videojogos 3D, XNA, Wiimote, Interacção Intuitiva.*

## 1. INTRODUÇÃO

Actualmente, tem havido uma grande evolução no que diz respeito à forma como os jogadores interagem com os videojogos. Esta interacção está a sofrer uma mudança dos controlos tradicionais para a utilização dos movimentos naturais do jogador.

Inicialmente, periféricos de entrada como joysticks, rato e teclado foram desenhados para aplicações 2D. Sendo assim, a sua utilização em aplicações 3D requer técnicas mais sofisticadas e maior treino na sua utilização. Apesar deste problema não se colocar no que diz respeito a jogadores experientes, é uma barreira para aqueles que não possuem grande experiência com videojogos. Por este motivo, actualmente existe um grande ênfase na criação de formas mais naturais de interagir com aplicações 3D, nomeadamente jogos de vídeo.

Este tipo de interfaces abre ainda a porta a outro tipo de utilizadores como, por exemplo, as pessoas mais idosas que não estão tão familiarizados com o tradicional rato e teclado. De acordo com um estudo de Ana Torres [Torres2008] os videojogos trazem benefícios aos idosos mas um dos principais problemas prende-se com a dificuldade de adaptação aos tradicionais dispositivos de interacção. Estes benefícios passam por várias áreas, como a reabilitação de movimentos motores, treino de memória, manutenção de boa forma, ensinamento de bons hábitos alimentares, entre outros.

Estes são alguns dos motivos que originaram a criação de comandos com detecção de movimentos na actual geração de consolas, com a Nintendo Wii, através do comando Wiimote a tomar o primeiro passo. A importância deste tipo de interfaces naturais está patente nas vendas da consola Nintendo Wii, que conseguiu destronar consolas com maior poderio gráfico, como a Xbox

360 da Microsoft e a Playstation 3 da Sony, que decidiram de início optar por comandos convencionais.

Com estas novas formas de interacção com videojogos, surge também a oportunidade de utilizar jogos para outros fins que não o puro entretenimento, surgindo aí os Serious Games. O termo Serious Game ainda não tem uma conotação aceite universalmente, mas é normalmente definido como um videojogo que imerge o jogador e simultaneamente contribui para alcançar um objectivo diferente do entretenimento, como a aprendizagem ou treino [Zielke2009]. A imersão é conseguida principalmente através do factor diversão, que é núcleo de qualquer jogo. Se um jogo não for divertido, torna-se aborrecido para o jogador, que acaba por abandonar o jogo ou, no caso dos Serious Games, que torna a aprendizagem ou treino pouco eficientes [Michael2006].

A diversão é portanto um factor muito importante, pois é através deste factor que o utilizador se abstrai do processo de aprendizagem ou treino, tornando assim estes processos mais apelativos e consequentemente mais eficientes.

Tem sido feita alguma investigação acerca de Serious Games que possam utilizar estas novas formas de interacção com jogos, nomeadamente através da detecção de movimentos. Alguma dessa investigação aborda temas como:

- Ensinar a escrever [Maxim2007].
- Reabilitação [Burke2009, DePonti2009].
- Manutenção de boa forma física [Hoysiemi2006].

Por exemplo, o WiiPlay permite que objectivos como a exercitação do jogador, ou até a reabilitação de movimentos motores dos braços possam ser concretizados, sem que o jogador se aperceba que o está a fazer. O

jogador fica entretido a jogar e abstrai-se do processo de reabilitação ou exercício.

O aspecto social dos videojogos tem vindo a tornar-se cada vez mais importante nos dias que correm. A maioria dos videojogos actualmente possuem algum tipo de jogo para mais do que um jogador, seja jogo cooperativo ou competitivo. O modo cooperativo oferece uma maior socialização aos jogadores, enquanto que o modo competitivo aumenta a competitividade no jogo.

Actualmente este aspecto é essencial e tem aplicações para além do aumento da diversão ou entretenimento, como o aumento da competitividade no trabalho, ou o aumento da capacidade de trabalho de equipa [Squire2005]. Estes aspectos são bastante importantes e segundo [Jansz2005], são as principais motivações para a participação de jogadores em LAN's. A competição foi apontada neste estudo como o aspecto mais importante para os jogadores, enquanto que a socialização foi considerada a segunda questão mais importante.

O sucesso dos jogos criados para as redes sociais exemplifica também a importância que o aspecto social tem para os videojogos, ou para o crescimento dos mesmos. Cada vez mais são utilizados jogos como um meio das pessoas socializarem, em vez de serem apenas um meio de entretenimento.

O WiiPlay possui também um modo de jogo cooperativo, em que dois jogadores cooperam para atingir um objectivo comum. Neste artigo é apresentado o jogo em si e são ainda descritos os testes efectuados para avaliar o potencial das novas formas de interagir com o jogo, em comparação com as tradicionais. No final são apresentadas algumas conclusões e trabalho futuro.

## 2. WIIPLAY

O WiiPlay é um jogo 3D de plataformas, criado através da plataforma XNA, visto na terceira pessoa através de uma câmara livre. O jogo possui dois níveis com temáticas diferentes, que possuem diversos obstáculos e inimigos que dificultam a progressão do jogador, de modo a proporcionar desafio. A principal novidade do jogo é a forma de interacção, que é conseguida através da detecção de movimentos com ajuda do Wiimote.

### 2.1 Wiimote

O comando Wiimote foi criado com o intuito de juntar a familiaridade de um simples comando de televisão com a tecnologia de detecção de movimentos, de modo a tornar os jogos acessíveis a uma maior audiência. Este comando inclui um microfone, um motor de vibração, quatro LED's3 azuis para feedback, um acelerómetro de 3 eixos, diversos botões (ver Figura 1) e também uma porta de expansão que permite ligar outros periféricos, tal como o comando Nunchuk. O Nunchuk contém a mesma tecnologia de detecção de movimentos que o Wiimote, mas inclui também um joystick analógico que ajuda no movimento da personagem controlada pelo jogador.

Através do uso da tecnologia Bluetooth, o Wiimote envia as acções do utilizador até uma distância de aproximadamente dez metros; Recorrendo a infravermelhos, pode funcionar também como um dispositivo apontador,

enviando sinal até cinco metros de distância. Podem ser ligados até quatro comandos (e respectivas expansões) em simultâneo, e podem também ser usados ambidextramente.

O Wiimote não requer nenhuma das características standard de autenticação ou encriptação Bluetooth, o que torna bastante fácil a sua utilização como periférico num computador normal. De modo a ligar o comando é necessário pressionar nos botões 1 e 2 (ver Figura 1), o que torna o comando detectável por outros dispositivos Bluetooth. Quando um botão é pressionado um pacote é enviado ao receptor através de um *HID input report* específico, que contém o estado actual de todos os botões. Este estado é também, incluído em todos os outros *input reports*. Os movimentos do controlador são detectados por um acelerómetro de 3 eixos lineares capazes de medir acelerações com um alcance de mais ou menos 3g com 10% de sensibilidade. O sensor usa um sistema de coordenadas ortogonal, baseado na regra da mão direita com o eixo X positivo para a esquerda, e o eixo Z positivo para cima quando o comando é segurado horizontalmente. Os inputs em cada eixo são digitalizados em *unsigned integers* de 8-bit. Em [Wiimote2009] pode ser encontrada a documentação que contém a lista completa de tipos de *reports* e os dados correspondentes transmitidos em cada *report*.



Figura 1- Ligação do Wiimote através de bluetooth.

### 2.2 XNA

O XNA é uma plataforma de criação de jogos baseada na Framework .NET e que utiliza a linguagem C#. Possui a vantagem de ser bastante fácil de utilizar, devido ao nível de abstracção que apresenta na sua estrutura e pelo facto de utilizar C# como linguagem de programação. O XNA utiliza o mesmo ambiente de desenvolvimento integrado (IDE) - XNA Game Studio Express - e o mesmo Framework - .NET - para o desenvolvimento de jogos tanto para Windows como para a consola Xbox 360, o que assegura um elevado grau de compatibilidade entre ambos. Apesar disso, existem algumas diferenças no que diz respeito à camada inferior. A Xbox 360 executa uma versão compacta da plataforma .NET, portanto existem algumas funções que estão disponíveis em Windows e que não executarão na Xbox 360.

O XNA suporta actualmente a utilização de vários periféricos, nomeadamente o comando Wiimote, que permite que seja utilizada uma interacção através dos movimentos do jogador, permitindo assim que sejam atingidos outros objectivos para além do puro entretenimento, como a manutenção da boa forma física [Hoysniemi2006], reabilitação [Burke2009], entre outras.

### 2.3 Aspectos do jogo

O jogo segue a abordagem clássica de jogos de plataformas, em que o jogador percorre os níveis de modo a coleccionar itens, enfrentando inimigos pelo caminho, assim como várias áreas de plataformas que tem de ultrapassar através da sua destreza e combinação de saltos.

O objectivo do jogo consiste em chegar ao fim de cada nível no menor tempo possível e com o maior número de itens recolhidos. Estes dados são guardados na forma de uma tabela de melhores resultados, que é apresentada ao jogador no fim de cada nível, ou no fim do jogo. Esta tabela pode também ser observada no menu inicial do jogo e tem como objectivo oferecer motivação através da competitividade. Assim o jogador tenta fazer melhores resultados que os outros participantes para ficar em primeiro lugar.

O jogo contém elementos muito importantes para a sua jogabilidade, que são os portais. Existem dois tipos de portais: portais de passagem de nível e portais de travessia dentro do nível. O jogador surge em cada nível através de um portal de passagem de nível e a única forma de avançar para o nível seguinte é passar por outro portal do mesmo tipo que é defendido por um inimigo bastante mais forte que os restantes (ver Figura 2). Os portais de travessia dentro do nível transportam o jogador para outras zonas do nível.



Figura 2 – Zona final - Jungle Level.

Os itens a apanhar estão espalhados por todo o cenário, obrigando o jogador muitas vezes a desviar-se da sua rota para os conseguir apanhar.

Existem vários inimigos, entre outras dificuldades a enfrentar de modo a poder avançar no jogo. Estas dificuldades adicionais passam por zonas de plataformas ou zonas de quedas de pedras que o jogador tem de evitar, por exemplo.

A única forma de avançar para o nível seguinte é derrotar um inimigo final de nível, que tem a particularidade de não sofrer danos provocados directamente pela personagem principal. Por este motivo, cabe ao jogador encontrar uma forma de o derrotar. Ao derrotar este

último adversário é revelado um portal de passagem de nível que permite que o jogador passe ao nível seguinte. Na Figura 3 pode ser visto um exemplo do segundo nível do jogo, intitulado Lava Level.



Figura 3 – Lava Level.

A personagem controlada pelo jogador possui três vidas e uma barra de vitalidade que indica a vitalidade actual. A personagem pode sofrer danos por quatro vezes. Ao sofrer mais do que esta quantidade de danos, esgota a sua vitalidade e perde uma vida, voltando ao início do nível. Quando se esgotam todas as vidas do jogador este perde, sendo o número de itens coleccionados e o tempo decorrido guardado numa tabela de melhores resultados, caso o jogador tenha obtido um bom resultado.

O número de itens coleccionados também é indicado na interface do jogo, de modo a que o jogador saiba imediatamente quantos itens já coleccionou (ver Figura 4).



Figura 4 – Interface do WiiPlay.

### 2.4 Jogabilidade

O WiiPlay utiliza o comando Wiimote da consola Nintendo Wii para conseguir uma jogabilidade através de detecção de movimentos. Deste modo, foi conseguido o tipo de interacção natural e intuitiva que se pretendia para o jogo, oferecendo simultaneamente uma interface familiar para os aficionados de videojogos, pois é um periférico bastante difundido em todo o mundo através da consola Nintendo Wii. Com este periférico foi alcançado também o objectivo de poder utilizar o jogo para outros propósitos que não entretenimento apenas. Ou seja, o jogador pode estar a jogar e ao mesmo tempo a fazer algum exercício pois são necessários movimentos cor-



porais para controlar a personagem principal, por exemplo.

O modo como a personagem se movimenta dentro dos cenários foi uma das grandes preocupações na criação deste jogo, pois o mais importante num jogo de plataformas é a sua jogabilidade. Para além disso, é essencial que um jogo seja divertido e ao mesmo tempo desafiante a nível da sua jogabilidade para criar motivação para o jogador.

No início foi utilizado apenas o comando Wiimote para o controlo da personagem principal, o que se revelou pouco intuitivo devido à quantidade de acções que são necessárias para um jogo deste género. Esta foi a razão principal que levou à decisão de aplicar uma das extensões do comando, o Nunchuk, de modo a separar as acções de movimentação das restantes acções.

A personagem é movimentada através do joystick analógico da extensão Nunchuk, pois este tipo de acção é bastante mais intuitiva desta forma do que através do uso de acelerómetros. Na Figura 5 pode ser visto um exemplo de como a movimentação do personagem é efectuada.



**Figura 5 - Movimentação da personagem através do joystick analógico do Nunchuk.**

O jogador tem também ao seu dispor um conjunto de acções de ataque (por exemplo pontapé ou murro), bem como, outras acções como rodar a câmara, saltar, correr ou simplesmente caminhar. Para rodar a câmara, o jogador roda o Wiimote na direcção que deseja rodar a câmara (ver Figura 6).



**Figura 6 - A rotação da câmara.**

Outro exemplo de um movimento feito como o Wiimote é a acção de salto. Este movimento é utilizado para ultrapassar as várias zonas de plataformas presentes no jogo e é feito através de um movimento intuitivo do comando de cima para baixo, como se pode verificar na Figura 7.



**Figura 7 – Salto do jogador sobre uma ponte.**

Esta forma de interacção permite também que o jogo possa ser jogado com base numa projecção e não apenas restringido ao monitor do computador. Este aspecto em particular pode ser bastante importante no futuro, essencialmente se pensarmos em novas formas de interagir com aplicações informáticas no geral, ou videojogos em particular. Por exemplo, permite desenvolver aplicações multi-utilizador mas que necessitam apenas de um único computador pois podem ser manipuladas num mesmo espaço físico com base numa simples projecção.

## 2.5 Jogo Cooperativo

O WiiPlay combina o jogo cooperativo com o jogo competitivo, na medida em que os jogadores trabalham em conjunto para atingir um fim comum (chegar ao fim de cada nível), mas cada um tenta coleccionar mais itens do que o outro, para obter uma melhor classificação do final de cada nível. Deste modo, o jogo permite a socialização entre os dois jogadores que têm de trabalhar em conjunto para chegar ao fim do jogo, mas simultaneamente competem entre si para obter melhores resultados individuais.

O modo de jogo para dois jogadores pode ser visto de duas formas distintas, o modo de ecrã partilhado e o modo de ecrã dividido. Se o segundo jogador estiver no campo de visão do primeiro e não estiver demasiado distante, então é utilizada apenas uma câmara e também o ecrã inteiro para mostrar os dois jogadores. Um exemplo do modo cooperativo com ecrã partilhado pode ser visto na Figura 8.



**Figura 8 - Jogo cooperativo com ecrã partilhado.**

Caso contrário, o jogo divide automaticamente o ecrã em dois e cada jogador é seguido por uma câmara distinta permitindo assim uma individualização das acções de cada jogador. No entanto, apesar de cada jogador ter uma visualização diferente, é possível verificar sempre a posição de ambos os jogadores através de um pequeno mapa. Este modo com ecrã dividido está ilustrado na Figura 9.



**Figura 9 – Jogo cooperativo com ecrã dividido.**

O mapa referido acima mostra todo o nível, ilustrando as zonas onde cada jogador se encontra no mesmo, através de uma pequena imagem que identifica a personagem controlada. A Figura 10 mostra uma representação mais detalhada deste pequeno mapa.

Existem algumas diferenças no que diz respeito à forma como a inteligência artificial dos inimigos presentes no jogo se adapta a um ou dois jogadores. No caso de existirem dois jogadores, a diferença reside no facto de os inimigos se focarem no jogador que está mais próximo, ou seja o jogador que apresenta mais perigo para os mesmos, a dado momento. No modo para um jogador apenas, o processo é bastante mais simples visto que os inimigos têm apenas um jogador para atacar.



**Figura 10 – Mapa indicador da posição dos jogadores**

Este modo cooperativo apenas permite que dois jogadores cooperem a nível local, pois actualmente apenas pode ser feito na mesma máquina. É também um objectivo futuro fazer alguns testes para verificar a qualidade deste tipo de modo para vários jogadores.

### 3. TESTES DE AVALIAÇÃO

Foram feitos alguns testes para avaliar o verdadeiro potencial da utilização da detecção de movimentos em jogos de vídeo, particularmente com o WiiPlay. Estes testes surgem na forma de um pequeno questionário que os participantes preencheram no fim de jogarem o WiiPlay com diversos dispositivos de interacção. Com estes testes pretendia-se perceber qual a forma mais intuitiva de jogar o WiiPlay. Estavam disponíveis as seguintes configurações: jogar com o Wiimote, jogar com o rato e teclado e jogar com o comando da Xbox 360. De referir que todos os participantes testaram as três configurações disponíveis.

#### 3.1 Estrutura do Questionário

O questionário é constituído por 11 questões e a sua estrutura está dividida em 3 partes distintas: identificação da região demográfica do participante, classificação da intuitividade do WiiPlay com as diferentes configurações e comentários gerais.

A identificação da região demográfica de cada participante tem como objectivo diferenciar o tipo de pessoa que participou no teste de avaliação. Deste modo, é possível comparar resultados de pessoas com experiência em videojogos com pessoas inexperientes. Cada participante foi classificado em termos de idade, área de trabalho, número de jogos que já jogou, tempo passado a jogar semanalmente e plataformas possuídas e preferidas, de modo a identificar quem tem experiência em jogos e se já utilizaram o Wiimote.

A classificação da intuitividade do WiiPlay com as diferentes configurações, consiste na classificação dos tipos de controlo utilizados (pouco intuitivo, intuitivo ou muito intuitivo) e na classificação da utilização do Wiimote para cada movimento disponível. Cada movimento disponível (movimentação da personagem, ataques, salto e rotação da câmara) é classificado como muito fraco, fraco, médio, bom ou muito bom.

São pedidos também neste questionário alguns comentários acerca do jogo WiiPlay e também acerca da utilização das várias configurações utilizados. Esta secção visa obter alguma informação que possa ter sido menosprezada durante a criação do questionário.

#### 3.2 Região Demográfica

Os testes foram efectuados num grupo de 20 pessoas com idades entre os 10 e os 50 anos, embora a maioria (75%) tivesse entre 20 e 30 anos. A maioria dos participantes (65%) já tinha jogado mais de 25 videojogos, enquanto que apenas 10% nunca tinha jogado um videojogo. A plataforma que a maioria dos participantes possui é o PC (85%), ficando ligeiramente à frente das consolas domésticas (65%). Apenas 15% dos participantes tinham experiência com controladores de detecção de movimentos como o Wiimote. De referir que vários utili-

zadores possuem mais do que uma plataforma de jogos, daí a soma das percentagens ser superior a 100%.

No que diz respeito ao tempo semanal gasto com videojogos, os valores são bastante variáveis:

- Nenhum tempo (20%).
- 2 horas semanais (35%).
- Entre 2 a 6 horas semanais (15%).
- Entre 6 e 12 horas semanais (15%).
- Mais de 12 horas semanais (15%).

De seguida são apresentados alguns resultados obtidos.

### 3.3 WiiPlay - Intuitividade

Nesta secção do questionário, foi pedido a cada participante que classifica-se cada controlador no que diz respeito à sua intuitividade.

No que diz respeito à utilização do Wiimote, o resultado do questionário foi unânime, tendo sido registadas apenas classificações positivas, indicando que o controlador é intuitivo/muito intuitivo. A maioria das classificações indica que este periférico é de facto muito intuitivo, tal como era de esperar. Este resultado pode ser verificado na Figura 11.

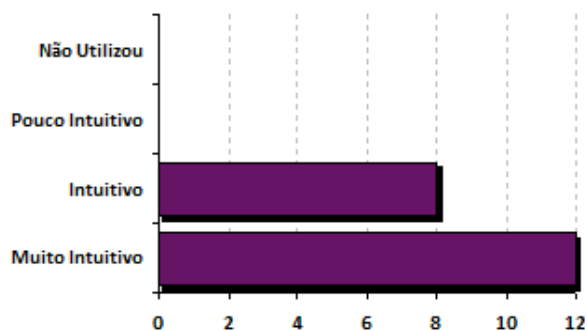


Figura 11 – Utilização do Wiimote

A utilização do rato e do teclado relevou-se pouco intuitiva, visto que a grande maioria dos utilizadores classificaram a sua utilização como pouco intuitiva. Apesar de a grande maioria dos participantes ter experiência em videojogos jogados no PC, a utilização do rato e teclado para este tipo de jogo foi a que menos agradou (ver Figura 12). O inquérito não incluía questões em relação a este facto, pelo que apenas podemos concluir que possivelmente este tipo de controladores não se adapta na perfeição a jogos com uma maior complexidade de movimentos, como é o caso dos jogos de plataformas. Este facto vem reforçar o argumento de que o rato e teclado, como ferramentas de interacção com videojogos, começam a estar desactualizados.

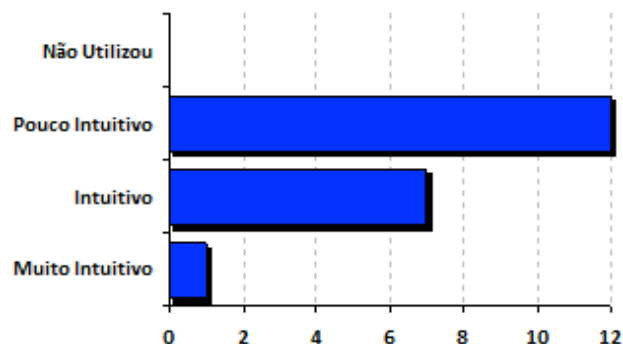


Figura 12 – Utilização do rato e teclado.

O controlador da Xbox 360 obteve também bons resultados, uma vez que é um controlador actual que é também bastante intuitivo. O resultado no entanto, fica aquém daquele obtido com o Wiimote. Na figura 13 é possível verificar que, neste caso, o controlador foi classificado na maioria dos testes como intuitivo.

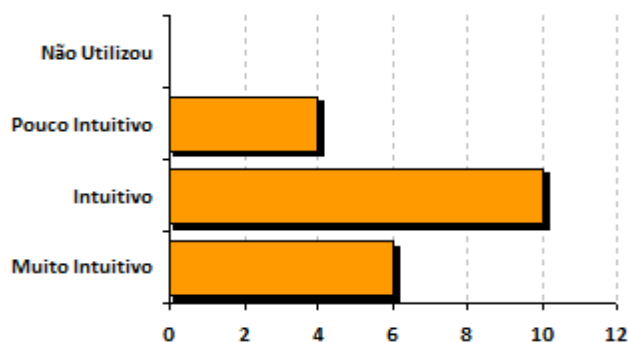


Figura 13- Utilização do controlador da Xbox 360

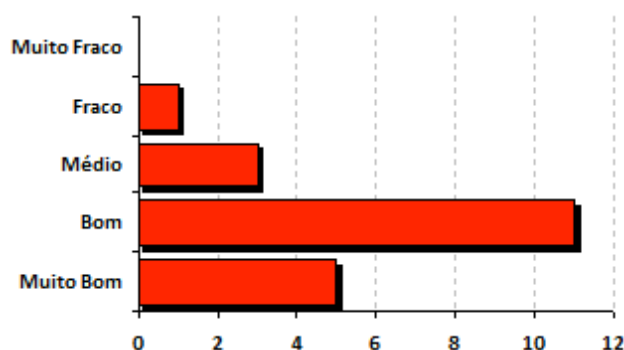
Segundo os participantes, o Wiimote foi o controlador mais intuitivo para jogar o WiiPlay. Este aspecto era esperado visto que toda a investigação consultada aponta para este resultado, confirmando assim a grande intuitividade oferecida por este periférico em particular e as futuras ferramentas de detecção de movimentos.

### 3.4 Classificação do Wiimote

O inquérito realizado pedia também que cada participante classificasse a utilização do Wiimote com maior pormenor, em relação a cada movimento utilizado no jogo. De seguida são apresentados os resultados desta classificação.

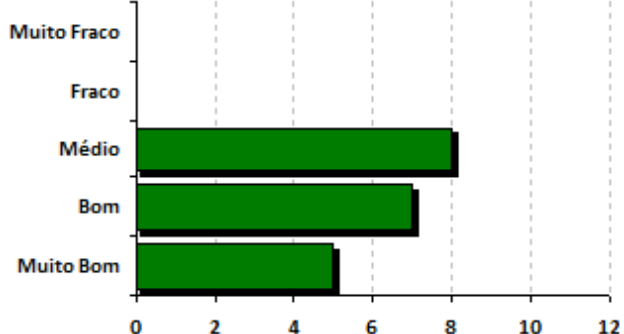
A movimentação da personagem obteve uma boa classificação, demonstrando assim que a utilização do Nunchuk vem ajudar à intuitividade do controlador, essencialmente devido ao facto do jogo possuir vários movimentos possíveis. Ainda assim, alguns participantes revelaram que a utilização desta extensão faz com que não sejam utilizados movimentos para todas as acções. A Figura 14 revela os resultados desta questão.





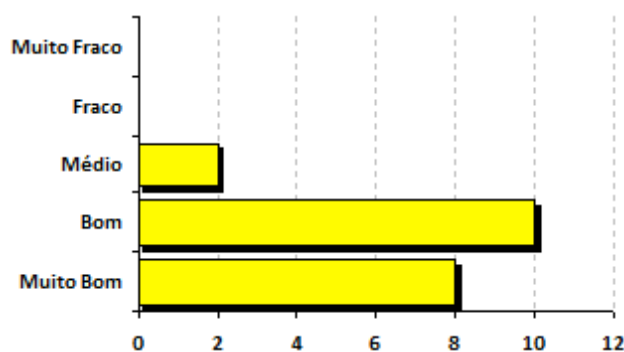
**Figura 14 – Classificação da movimentação da personagem com o Nunchuk.**

Os ataques obtiveram classificações diversas, mas nenhum utilizador classificou os mesmos como sendo fracos (ver Figura 15 para mais detalhes). Alguns utilizadores referiram no entanto que poderia haver mais tipos de ataques, como ataques especiais que pudessem ser desencadeados através de determinadas movimentações do Wiimote. Assim, seria tirado ainda mais partido das propriedades do acelerómetro do Wiimote.



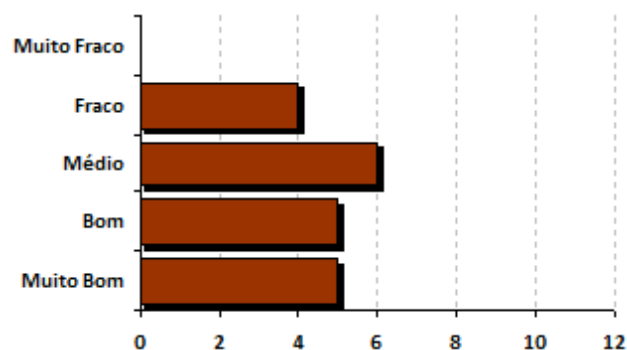
**Figura 15 – Classificação dos ataques.**

A classificação do salto foi o movimento que obteve melhores resultados. Este resultado explica-se por ser um movimento que resulta da utilização exclusiva do acelerómetro, num movimento de mão semelhante ao que o utilizador quer que a personagem faça. Ou seja, o jogador faz um movimento de baixo para cima com o comando, que resulta num movimento de baixo para cima (salto) da personagem. Ainda assim, alguns dos participantes indicaram que este movimento deveria ser mais sensível, de modo a tornar a interacção ainda mais intuitiva. A Figura 16 mostra os resultados deste teste.



**Figura 16 – Classificação do salto**

Em relação ao movimento de rotação da câmara, foi o movimento que obteve classificações mais distintas, desde fraco, a médio, a bom e a muito bom (ver Figura 17). Apesar de algumas pessoas acharem ótimo, outras acharam demasiado difícil o movimento de rotação do pulso para rodar a câmara. Para além deste facto, alguns participantes revelaram que devia ser possível movimentar a câmara também no eixo Y.



**Figura 17 – Classificação da rotação da câmara.**

Para além destas classificações, os utilizadores fizeram algumas observações e comentários acerca da experiência em geral. Esses comentários são apresentados na próxima secção.

### 3.5 Comentários Gerais acerca do WiiPlay

Os participantes no teste fizeram bastantes comentários acerca da experiência geral, onde revelaram algumas informações interessantes. No geral, os participantes classificaram o jogo de bastante divertido e apelativo. Alguns afirmaram que o Wiimote permite um maior grau de liberdade e que é bastante mais intuitivo para quem não tem experiência. No entanto, vários participantes apontaram como uma melhoria a existência de um maior leque de movimentos possíveis, para conferir maior diversidade ao jogo com o objectivo de o tornar ainda mais divertido.

Um comentário bastante interessante referido indicava que apesar de o jogo ser bastante engraçado, poderia ter alguma vertente formativa. Deste modo, poderia ter um propósito diferente da diversão apenas e contribuir para o crescimento pessoal a nível de aprendizagem tanto em jovens como em adultos.

Outros participantes questionaram o facto de no teste não estar nenhum modo multi-jogador disponível e realçaram que consideram esse aspecto muito importante. Apesar desse modo estar disponível, não foi ainda utilizado para classificar a interacção. No entanto, esse aspecto é cada vez mais relevante ao ponto de desencadear interrogações por parte de alguns participantes pelo facto de o mesmo aparentemente não estar disponível.

## 4. CONCLUSÕES

Foi apresentado o WiiPlay, um jogo de plataformas em três dimensões, visto na terceira pessoa e cuja jogabilidade é conseguida através de movimentos naturais do jogador com o auxílio do comando Wiimote.

Foi efectuada também uma avaliação da interacção com o jogo com base num questionário, de forma a classificar a utilização dos diversos periféricos de interacção disponíveis para o WiiPlay, incluindo periféricos tradicionais e um periférico que permite a detecção de movimentos (Wiimote). Os resultados obtidos mostraram que a utilização do Wiimote é bastante mais intuitiva do que a utilização dos restantes periféricos mais tradicionais. Os resultados mostram ainda uma preferência por parte dos participantes na utilização das propriedades dos acelerómetros do Wiimote, ao invés da utilização dos seus botões.

Conclui-se também que o aspecto social dos videojogos é cada vez mais importante e frequente, sendo este um aspecto bastante promissor nesta área.

## 5. TRABALHO FUTURO

Apesar da utilização do Wiimote se ter relevado bastante produtiva e com resultados bons no que diz respeito à intuitividade, existem já novas tecnologias promissoras para detectar os movimentos do jogador. Por esse motivo, o trabalho futuro passará obrigatoriamente pela avaliação dessas tecnologias.

O aspecto social do jogo deverá também ser explorado em maior profundidade, com o teste do modo de multi-jogador criado e possivelmente com a expansão para um modo cooperativo online.

Através do feedback recebido nos inquéritos percebeu-se que a criação de um videojogo com um propósito explícito de aprendizagem ou treino poderá ter mais impacto como Serious Game, do que apenas um jogo facilmente adaptável para o propósito de um Serious Game. Assim, trabalhos futuros poderão passar pela criação de Serious Games explícitos, mantendo no entanto intacto o factor de diversão para tornar o jogo numa ferramenta eficiente de aprendizagem ou treino.

## 6. AGRADECIMENTOS

Ao Instituto de Telecomunicações (IT) que suportou os equipamentos necessários e concedeu uma bolsa ao primeiro autor para a realização deste trabalho.

## REFERÊNCIAS

[Burke2009] J. W. Burke, M. D. J. McNeill, D. K. Charles, P. J. Morrow, J. H. Crosbie and S. M. McDonough.

Serious games for upper limb rehabilitation following stroke. *Proceedings of the 2009 Conference in Games and Virtual Worlds for Serious Applications*, pp.103-110, 2009.

[Deponti2009] Dario Deponti, Dario Maggiorini and Claudio E. Palazzi. DroidGlove: An Android-Based Application for Wrist Rehabilitation. *ICUMT 2009*, pp.1-7, 2009.

[Hoysniemi2006] J. Hoysniemi. International survey on the dance revolution game. *ACM Computers in Entertainment*, 4(2), 2006.

[Jansz2005] Jeronen Jansz, Lonneke Martens. Gaming at a LAN event: the social context of playing video games. *new media & society*. SAGE Publications. 7(3):333-355, 2005.

[Maxim2007] Bruce R. Maxim, Niles V. Patel, Nicholas D. Martineau, and Mark Schwartz. Work In Progress - Learning via Gaming: An Immersive Environment for Teaching Kids Handwriting. *37th ASEE/IEEE Frontiers in Education Conference*. pp.T1B-3 - T1B-4, 2007.

[Michael2006] D. Michael and S. Chen. Serious games: Games that educate, train, and inform. *Course Technology Boston*, MA, 2006.

[Squire2005] K. Squire and C. Steinkuehler. Meet the gamers. *Library Journal*, 130(7):38-42, 2005.

[Torres2008] Ana Torres. Cognitive effects of video-games on older people. *ZON DIGITAL GAMES*, 21-27, 2008.

[Wiimote2009] Wiimote low level details. 2009.  
<http://www.wiili.org/index.php/wiimote>

[Zielke2009] Marjorie A. Zielke, Monica J. Evans, Frank Dufour, Timothy V. Christopher, Jumanne K. Donahue, Phillip Johnson, Erin B. Jennings, Brent S. Friedman, Phonesury L. Ounekeo, and Ricardo Flores. Serious Games for Immersive Cultural Training: Creating a Living World. *IEEE Computer Graphics and Applications*, 29(2):49-60, 2009.

# Augmented Reality Without Markers

Luis F. A. Silva      Frutuoso G. M. Silva

Instituto de Telecomunicações

Universidade da Beira Interior

Rua Marquês D'Ávila e Bolama, Covilhã, Portugal

m2086@ubi.pt, fsilva@di.ubi.pt

## Abstract

*This paper introduces an augmented reality application for web that works without markers. We propose a solution to manipulate virtual objects without using markers, i.e. we replaced the physical marker by user face. For this we had to develop also a method to manipulate the virtual models (e.g. to rotate models) based on user gesture allowing doing the same as the marker. Finally we present a simple augmented web game showing our approach without the use of markers.*

## Keywords

*Augmented reality without markers, Augmented web game, Gestural interface*

## 1. INTRODUCTION

The constant growth of the Internet, where a variety of new ideas and information are published every day, make it in an excellent media to present information. Recently, Augmented Reality (AR) has been used also as a way to publish information in Internet. But we can see that the augmented reality had a weak growth until 2001 [Azuma 01], but recently is visible a large adherence to AR, particularly in marketing and dissemination of products and concepts. It's worth highlighting as examples the marketing of the latest movies, Star Trek<sup>®</sup><sup>1</sup>, Avatar<sup>®</sup> and among others by AR. Sales are another area where the onset of AR has been noticed, where brands such Citroën<sup>®</sup> (C3 Picasso) or Nissan<sup>®</sup> (Cube) are using AR to advertise their models.

According to Azuma [Azuma 01], we have an augmented reality application when we add virtual objects to a real environment, i.e. the objects created by computer (i.e. virtual objects) are inserted in a real environment where the real environment is normally captured by a camera. But the augmentation is conventionally in real time and allowing user interaction. Because what distinguishes AR from a simple 2D/3D scene placed over an image, captured by a camera, is the interactivity. In other words, in AR the users have the power to interact in real environment with virtual objects, thus giving a perception that these objects are in their world.

In general, in most applications of AR the objects are added to the real environment according to the marker captured by digital camera [Liarokapis 04b] or by geographic coordinates provided by GPS [Narzt 06], between other

methods. But most AR applications are using markers to positioning and orienting the virtual objects.

Recently, Zhou *et al.* [Zhou 08] presented a study about the trends in augmented reality, tracking, interaction and display, in last ten years. In this study, the authors say that hand gesture recognition is one of the most natural ways to interact with an AR environment but this in the context of tangible interfaces. In the context of gestural interfaces, they identify only one method based on finger tracker. Unfortunately, in this work we not find any reference to augmented reality applications for web browser. However, Liarokapis *et al.* [Liarokapis 04a] presented an educational application that allows users to interact with 3D Web content using virtual and augmented reality.

In this paper we describe a web augmented reality application that works without the use of markers. This application have been developed to web, and it is available for those that have a web browser and a webcam. Section 2 describes our approach that works without markers to manipulate the virtual objects. In section 3 is presented the augmented web game developed to show the potentialities of our approach. Lastly, section 4 presents some conclusion and future work.

## 2 AR without Markers

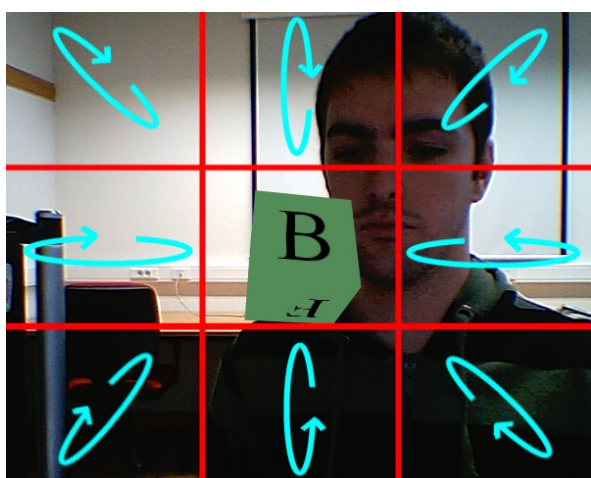
One of the objectives of our work is simplify the interaction, for that reason we have been removed markers, being the interaction made only with the body of the user (hands and head). The alternative chosen as reference point for the model was the use of the face user. Thus it is possible to put the virtual model in the scene according to the position of the face user. However, this alternative not al-

<sup>1</sup>see [www.experience-the-enterprise.com](http://www.experience-the-enterprise.com)

allows to rotate the model as we can do with markers. Hence we divided the screen in several areas allowing the rotation of the model based on motion detection for each area (as shown in Figure 1).

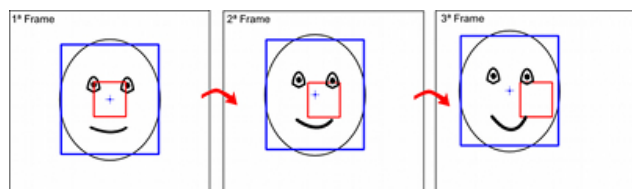
Then application makes the following calculus based in face and motion detection for each frame:

- Calculation of  $z$  co-ordinate is based on the user's face, i.e. the distance that he/she is from the webcam;
- Calculation of  $x$  and  $y$  is based on the position of the face user in the environment (i.e. similarly with the markers);
- The orientation of the object is made by detecting motion of the user in different areas of the screen.



**Figure 1. Interaction areas for rotating the object based on motion detection.**

In short, the face of the user will be used as reference point to positioning the objects in virtual environment and the motion detection to interact with virtual object (e.g. to rotate). To prevents a conflict between the functions of face detection and movement detection, it was considered a priority rule to establish a functional hierarchy that had success. So, first is considered the face detection phase and if this function returns a new detection, the motion capture is aborted. If the face detection is not significantly changed it's taken into account the motion capture phase. Normally,



**Figure 2. Example of the threshold in a sequence of frames.**

the face is presented in all frames but it is only considered

a new detection when it is in a different position from last detection. Thus was created a threshold to the position of the face, when this threshold is exceeded it is considered a new detection. This threshold is calculated based on the size of the face and the screen resolution (see Figure 2).



**Figure 3. Rotating the object**

The areas of interaction are pre-defined. So if is detected movement in right area of the screen, the virtual object will perform a rotation in  $y$ -axis (rotation to the right). The rotation in  $x$ -axis is activated if is detected movement in top or bottom area of the screen. It is also possible make a rotation simultaneously on two axes  $x$  and  $y$ , which will be triggered if there is a movement in one corner of the screen. Note that the motion detection is restricted to small



areas, i.e. it is not allowed in large areas. This prevents that large movements in area are considered, because most of these situations are not orders of rotation. Only the movements detected for these small areas will be considered as instructions for the application (e.g. a simple movement of the fingers in these areas). Figure 3 shows some examples of the use of the application. In this example the virtual object is a 3D logo of the Institute of Telecommunications (IT) and the user can rotate it based on movements in pre-defined areas.

### 3 Augmented Web Game

To show the potentialities of our approach we develop a web game using augmented reality without use of markers. As explained in previous section we use the user face as reference to put the virtual model in the scene and its manipulation is based on gesture of the user.

The main idea of the game is to find the letters of a word. For that, we have a virtual cube, which has a single character on each face, and the user has to rotate the cube to find the characters of the word (see Figure 4). The word is presented below the cube using several plans, one for each character of the word to find. These plans are initialised with exclamation character (i.e. '?'), which is replaced for the correct character when it appears in front face of the cube. Thus the user need to rotate the cube to find the characters of the word. The rotation of the cube is available in two axis,  $x$  and  $y$ , and it can be activated by the user movements only in four special areas, namely in upper, bottom, left and right area of the image. Note that it is possible use the eight areas of interaction with the model, as described in previous section, but we use only four to simplify the game (i.e. using only four areas is more easy to maintain the front face of the cube aligned with user).

When the user activates a rotation of the cube, the game generates a visual feedback to help him. The feedback helps user to understand which type of movement was realised. This feedback is created using a set of arrows that are moving in the direction of the rotation, as shown in the second and third pictures of Figure 4. The game has a timer clock and shows the number of words that are available yet. This information is displayed above the cube, as shown in first picture of Figure 4.

The position and size of the cube depends of the position and size of the face of the user, respectively as explained in previous section. However, when the face of the user disappears of the image, the cube stays in the last position of it (as shown in Figure 4).

The game uses a XML configuration file with the words of the game. Thus, it is possible to change the words of the game by changing this XML file. This simple game shows that our approach is viable and can be adopted in other applications on web.

### 4 Conclusion and future work

The generality of augmented reality applications are dependent of markers to run. But we presented a new approach to run without markers. Our augmented web game

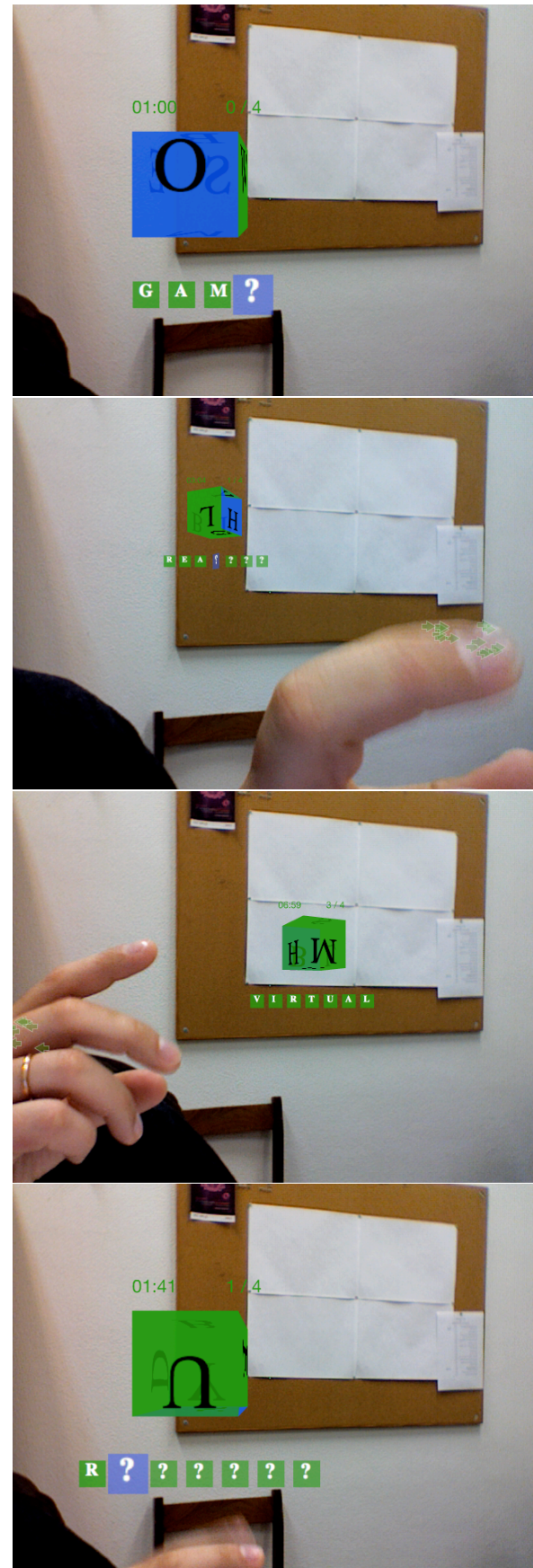


Figure 4. Augmented web game

showed that this new way of interaction without markers is useful and can be a simple solution for augmented reality applications in web browser. Therefore it is possible have augmented reality applications without physical markers with similar functionalities to applications that use physical markers. Note that the application was developed in Actionscript and compiled for Adobe Flash. Therefore the application has the ability to be used via web browser by any user that have the Flash Player plugin.

In future, we want to make an evaluation of the new way of interaction without markers with a group of users. Besides, we want to see how this type of applications will run using a web API for creating interactive 3D applications in browser, as WebGL [WebGL ] from Khronos Group.

### Acknowledgments

The first author would like to thank the grant given by Instituto de Telecomunicações to support his work.

### References

- [Azuma 01] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *IEEE Computers & Graphics IEEE*, 21(6):34–47, 2001.
- [Liarokapis 04a] Fotis Liarokapis, Nikolaos Mourkoussis, Martin White, Joe Darcy, Maria Sifniotis, Panos Petridis, Anirban Basu, and Paul F. Lister. Web3d and augmented reality to support engineering education. *World Transactions on Engineering and Technology Education*, 3(1):11–14, 2004.
- [Liarokapis 04b] Fotis Liarokapis, Martin White, and Paul Lister. Augmented reality interface toolkit. In *IEEE Proceedings of the Eighth International Conference on Information Visualisation*, pages 761–767, 2004.
- [Narzt 06] Wolfgang Narzt, Gustav Pomberger, Alois Ferscha, Dieter Kolb, Reiner Müller, Jan Wieghardt, Horst Hörtnner, and Christopher Lindinger. Augmented reality navigation systems. *Univers. Access Inf. Soc.*, 4(3):177–187, 2006.
- [WebGL ] WebGL. Webgl - opengl es 2.0 for the web. Khronos Group, <http://www.khronos.org/webgl/>.
- [Zhou 08] F. Zhou, H. B. L. Duh, and M. Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. In *7th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2008)*, pages 193–202, 2008.



# 2D Game Editor On-line

## A Cloud Computing Perspective on Game Edition

Bruno Miguel Cardoso  
CITI-DI/FCT/UNL,  
Departamento de Informática  
Universidade de Évora  
Rua Romão Ramalho 59  
7000-671 ÉVORA, PORTUGAL  
+351 266 745373  
b.m.pinto.cardoso@gmail.com

Teresa Romão  
CITI, Departamento de Informática  
Faculdade de Ciências e Tecnologia  
Universidade Nova de Lisboa  
2829-516 Caparica, PORTUGAL  
+351 212948536  
tir@di.fct.unl.pt

### Abstract

*The present article aims to describe the efforts taken in order to create a generic computing solution for the most recurrent problems, usually found in the production of two dimensional, sprite-based videogames, running on mobile platforms. The developed system is a web application that fits within the scope of the recent cloud-computing paradigm and, therefore, enjoys all of its advantages in terms of accessibility, application maintainability and scalability. In addition to the functional issues, the system is also explained in terms of its software architecture, since it was planned and implemented in the perspective of attaining an easy to maintain application that is both scalable and adaptable. Additionally, an algorithm is also proposed, seeking to find an optimized solution to the problem of space distribution for several rectangular areas, with no overlapping and no dimensional restrictions, neither on the final arrangement nor on the arranged areas.*

### Keywords

*Game edition, web based solution, scalability, software engineering, collaborative work.*

## 1. INTRODUCTION

Facing a growing activity in the last thirty years, the market of electronic entertainment – of which videogames are the most preeminent representatives – moved, just in 2007, over 41 billion dollars across the whole world. This value far surpasses the ones presented, for the same period, by the movie industry [Alves2008]. However, the production of videogames that match current market standards can be a daunting task, since it may require expertise in several artistic and technological fields, especially when the new title is intended to run on advanced consoles or personal computers.

Still, there is a particular kind of videogame whose production is much less demanding and still enjoys a wide acceptance by the consumers: the casual games. Being much smaller (in every aspect) than a traditional title, this kind of game also has much simpler rules and, thus, requires a far less dedicated gamer profile. This contributes, of course, to widening the range of its potential consumers – in fact, they have a possible user in just about anyone who wishes to spend some time entertaining themselves. Furthermore, these already widened use scenarios can still be extended if the game is targeted to run on mobile platforms. That way, a possible use context can be identified in every place where such a platform is able to operate. Summarizing this point, not

only a casual game can be used by almost anyone, they can also be used almost everywhere.

However, although being simpler to create and distribute, the production of a new casual title may still require the same specialized skills that the more traditional and advanced videogames do – only perhaps not so deeply. Even though it's possible for a single professional to have all of the necessary skills, such is not, indeed, a realistic premise. It's far more common – and, perhaps, even preferable – to invest in multidisciplinary and enlist a team of various professionals, skilled in each of the required fields of knowledge. Clearly, this will bring an extra effort to be made, in order to combine their distinct works (e.g. art, written texts, sound) into really useful, programmatically accessible resources.

Also, there is a peculiarity in general videogame development that may promote more constraints to its production. In a nutshell, the game has a concept that must not be lost during its development. This concept defines the user experience, general gameplay and provides a detailed overview on how the game should perform once implemented. Being such a comprehensive and abstract idea, the game concept may also include subjective impressions and other not so easily transmittable aspects. The responsibility of maintaining the development processes oriented to this concept

usually falls upon one professional, the game designer, who must be, before anything else, a creative and resourceful person with some technological and artistic perceptiveness. Then, in order to create a game that is true to its concept, another of the game designer's duties refers to coordinating those heterogeneous teams, where its members may have their own professionally-specific technological languages (that can even be, to some extent, quite hermetic). Since a significant part of the designer's tasks consist in transmitting the game concept's ideas to the development team, it follows that the threat posed by the occurrence of communication misunderstandings can be exceptionally dangerous to the project itself.

To reduce the possibility of human error emerging from the natural language's intrinsic subjectivity, a recurring solution is frequently adopted: the usage of game editing applications. These systems offer, through interactive and friendly user interfaces, the possibility to parameterize certain game properties that further tune the game player's experience. Such kind of solution can be found both in amateur and professional game productions, since they all have the same basic requirements. However, mainly due to resource availability, there are notable differences between the two. While industrial editors use to be larger, internally developed, desktop applications that provide support to a wide range of tasks, the editors used in amateur productions tend to be smaller, single-tasked applications that are developed by other amateur producers, who, then, distribute them freely over the web. Being larger applications, the industrial editors also have more effort applied to their development and, as so, they are not lightly replaced or discarded. Instead, every time a new and sufficiently valid production requirement arises, they are likely to be subject of maintenance interventions, with the purpose of adding, altering or removing existing functionalities. Such maintenance procedures can be achieved by creating software patches or, more often, by tampering directly with the editor's source code. Either way, the needs that motivate these "ad-hoc" interventions can have some degree of urgency which may force the maintainers to disregard the (eventually) existing code architecture and methodologies. This will end up in posing difficulties to future maintenance tasks.

Besides from programmatic disadvantages that this "ad-hoc" development methodology is likely to impart, the application's user interface also gets prejudiced. In fact, specific widgets that once have been assigned to operate on specific areas of an interface, may now find themselves competing with lately developed controls that offer whole new functionalities. Such occurrences are hardly positive for already experienced users, even in the unlikely case that the developers had the time to carefully plan the alteration, and may ultimately lead to software abandon [Cardoso08].

On the other hand, because the free editors used in amateur productions are independent applications, the production lines that use them will have to include efforts

to merge their different outputs and to make sure they integrate, as intended, into the developing game.

There are several such applications on the web, each one being specially oriented to assist in a specific task. Particularly for sprite based videogames, a most useful functionality relates to sprite packing. This means aggregating the images that compose a sprite's animations, the frames, into new, bigger images, the atlases, in an optimized – but not necessarily optimal – way. In fact, the frames usually come from the artists who create them arranged in larger images, known as strips. These strips are rather raw resources and, seeing as they have a lot of unused space between the frames, they are too much "unwieldy" to be used in the final product. Indeed, since these large images would have to be loaded into memory, using them directly would inevitably result in increasing a game's memory requirements. So, a sprite packer typically operates by receiving the already cut frames, making an arrangement out of them and exporting both the produced atlas and the necessary metadata references that allow external agents (e.g. the game engine) to univocally refer to a specific frame within the atlases larger scope.

However, these free sprite packers fall short of being truly supportive applications, because they do just that - pack the frames passed in as input - while neglecting that there are other frame related tasks to accomplish (e.g. cut them out of the strips or parameterize the sequences that are, in fact, the animations themselves). Some examples of these free editors, programmed in various languages, would be the Packer<sup>1</sup>, from the 2D Slick Game Library, and the Image Packer<sup>2</sup>. Even though these applications are freely distributed over the Internet, they are not true web applications; instead, they are small programs that must be downloaded and locally executed on the client machine. As previously stated, none of them allows for anything else than packing the frames and some do not even have anything but rather rudimentary atlas composing algorithms.

With the intention of simplifying the implementation of such algorithms, the developers often place on the users the responsibility of making some very important decisions, like the final dimensions of the produced atlas. This may have significant impacts on the application's real utility and even contribute to degrade its performance since not only direct user intervention will be required, but also it makes way for poor dimensional choices that may imply large amounts of unused area.

## 2. THE 2D GAME EDITOR ON-LINE

The system proposed in this paper, named 2D Game Editor On-line (2DGE-O), consists of a web application intended to provide support for the most recurrent tasks found in two dimensional, sprite based videogame production. These tasks include: strip loading; selective cutting of the frames embedded on a loaded strip; aligning the sequences of frames that compose a sprite's

---

<sup>1</sup> <http://slick.cokeandcode.com/>

<sup>2</sup> <http://www.pascalgamedevelopment.com/forum/index.php>

animations and exporting the manipulated information to standard formats, such as XML and .png atlases. To further enhance the exporting process, an algorithm has been perfected to automatically create optimized atlases. One of the system's key features lies on its strong end user orientation and, to this end, much attention was devoted to its user interface. The system is implemented with a mix of technologies, like PHP and MySQL on the server side and, on the client's end, DHTML, Java Applets and AJAX for asynchronous communications with the server. Fig. 1 shows the frame cutting Applet (on the right, below the "Strip blaze" label), embedded on the system's general web interface.

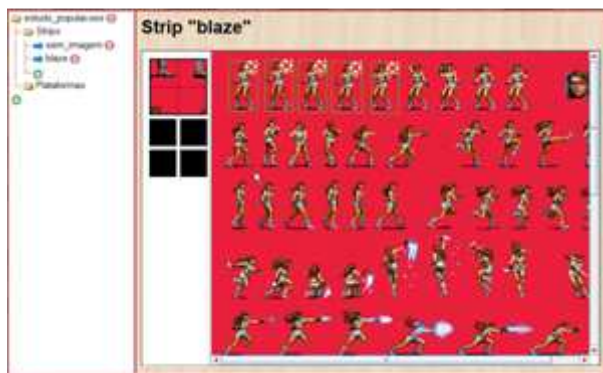


Fig. 1- 2D Game Editor On-line interface overview.

An essential difference that conceptually distinguishes this solution from the ones described in section 1 lies in its way of integrating the multitude of tasks, necessary to produce a videogame, into a single concept, the project. This integrative perspective closely resembles the one that is traditionally found on industrial game editors.

### 2.1 Game Edition in Cloud Computing Paradigm

The 2DGE-O is a web application that, according to the NIST's definition of Cloud Computing [Mell09], adheres to the Software as a Service Model. This means that the user does not control or manage the underlying cloud infrastructure, and it offers all of its functionalities with the sole requisition of a thin client interface - in this case, a regular browser and some of its common resources, like the JVM<sup>3</sup>, cookies and JavaScript enabled. This allows for a game development team to be geographically dispersed and still be able to collaborate on the same game projects. In another perspective, that of the developer's, the main advantage that emanate from basing such a system on the web, is the possibility of focusing the maintenance efforts on just one copy of the application - precisely the one that is installed on the production server. Any modifications performed on this server's deployed 2DGE-O instance will instantly be made available for all of its users.

### 2.2 The System's Internal Architecture

Noting the enormous product variety on modern videogame's market, one can infer that is impracticable to create an editor that attends to all of this industry's requirements. Then, to cope with this impossibility,

another objective of the 2DGE-O is to provide the basis for future functional module development. In this sense, there was a necessity for programming methodologies that contributed to maintaining the code as human readable as possible, while defining clearly the procedures to follow when creating new components. Such sort of problem is not new in the scope of software engineering. In fact, these requirements are so commonly encountered, that they have given origin to more than one solution meant to solve them, generically referred to as "architectural patterns". This concept is defined as an *"idea that was useful in a specific practical context and that is likely to prove useful in others"* [Fowler96].

There are several architectural patterns that enforce code methodologies and, at the same time, keep a close attention to the user interface and its respective mechanics. The most representative would be the Model-View-Controller (MVC). In spite of having been first proposed to be applied on *SmallTalk* environments [Krasner88], this pattern was the most adequate to the 2DGE-O characteristics. The MVC imposes, as stated in [Buschmann96], the separation of an application's code in triads made of three functional components, according to the responsibilities they assume in the system's functionalities: the Model is responsible for the data and provides the methods to operate on it; the View renders the corresponding model's information to the user and, lastly, the Controller reacts to and makes decisions based on user input. The user interface (or parts of it) is formed by View-Controller pairs. Such pairs can be connected directly to another triad's Model, if the necessary methods are already implemented in it. This way, in order to create a new functional module and expect it to integrate smoothly into the already existing interface, the programmer must implement only - in the best case - a new View-Controller pair and associate it with an already existing Model. Of course, if there is no Model that already performs the actions required by the new pair, then a new one will have to be implemented. Such well defined procedures contribute greatly to maintain the 2DGE-O's code legible and organized after such alterations are completed.

### 2.3 The Atlas Composing Algorithm

The usage of texture atlases in games can bring many advantages to their performance and this is especially true when the game is intended to run on low resourced platforms, such as mobile phones. They allow saving memory that would, otherwise, have to be allocated to very consuming tasks, like image swapping. Since a game might have several sprites rendered simultaneously on screen, these operations can turn into considerable nuisances.

An algorithm, proposed by [Scott], configures an interesting option to solve the problem of packing a number of rectangular areas (the original proposal considers these as being lightmaps) into a fairly optimized image. To this end, it represents a two-dimensional empty area (the atlas) as a binary tree that registers growth whenever new images are added and

<sup>3</sup> Java Virtual Machine.

arranged. Nonetheless, there are some issues with the described approach that drastically detracts from it, namely, it assumes the final atlas dimensions to be known *a priori* or, at least, an upper bound for them. This is not a very practical solution, since the users will end up searching for a satisfactory result through a *generate and test* (or *trial and error*) methodology. To avoid this, it's necessary to find a way to provide more autonomy to the algorithm.

We found a solution by supplying an empty rectangular area of unlimited dimensions and define it as the base upon which the arrangement will be made. Inside it, another rectangular boundary will be kept in memory, first with dimensions set to 0 and then, as more frames are added, it will be iteratively enlarged to accommodate them. This area is referred to as the "frontier". The key feature of the algorithm lies in an heuristic approach that provides guidance every time this frontier needs to expand. Many approaches were attempted, but the most satisfactory results have been achieved when using a special scalar measure,  $R$ , that, for any rectangular area,  $R = \frac{\text{shorter side}}{\text{longer side}}$ . It's trivial to deduce that, for any given rectangle,  $0 < R \leq 1$ , and that  $R = 1$  only when the area is a perfect square. The algorithm starts by calculating the average  $R$  for the whole population of frames that will be arranged. Then, it proceeds to successively inserting the frames and, whenever one is encountered that cannot be arranged inside the current frontier, the algorithm will expand its borders, choosing the resize that grants the closer proximity between the frontier's own  $R$  and the frame population's average  $R$ .

### 3. USER TESTS

The user tests were performed in two sessions, with the same testing group composed of five users, whose average age was about 35 years and all have worked with IT for a period of 2 to 14 years. One of them was an amateur game designer. These sessions followed a previously planned script that included the main tasks an editor is required to perform, such as uploading strips, cut its frames and building a few animations with them.

In the appointed tasks, the user's feedback was mostly positive, consistent and very focused on the functionalities offered by the application. The user interface revealed itself both as easy to use and easy to learn. Apart from the positive comments about the system's functionalities and on how they were implemented, there was a significant interest on the space allocation algorithm and its potential application to problems outside the context of videogame production (a possible example would be the reduction of website images download time).

There were, however, some invaluable remarks made by the users that contributed to further improve the system's interaction with the end-users. As an example, one of the registered observations suggested further refinement of the included zooming tool, so that it may provide more in-depth support to the tasks of frame selection and cutting.

### 4. CONCLUSIONS AND FUTURE WORK

The 2D Game Editor On-line results of an innovative approach to a recurrent problem, i.e. the cloud-computing paradigm's answer to the game edition's requirements. Put simply, the system is a web application, accessed through common browsers, which allows for remote collaboration between geographically dispersed teams. Also, because the source code is solely located on the server, the client-server model eases the efforts associated to application maintenance. This way, when necessary to alter the current version, the maintainers can intervene directly on the server, while keeping the clients unaware, or unaffected, by the operations.

Because of the MVC pattern applied to the system's architecture, the system ends up being highly scalable, functionally flexible and maintainable.

The user's feedback during the tests was very positive. However, despite the fact that the 2DGE-O gives full support to sprite manipulation tasks, there are a number of improvements to be done on the already implemented functionalities, which will greatly add to the system's usability and performance. As an example, the atlas composition algorithm could be further optimized if allowed to operate with rotation. This way, before deciding to expand the current atlas's frontiers, the algorithm should attempt to reinsert the same frame, after a 90° rotation. As future work, there will also be implemented a undo/redo functionality that allows users to revert their work to previous states, as well as the user suggested refinements to the zooming tool (referenced on section 3). Additional user test to evaluate the system's usability and performance will also be performed.

### 5. REFERENCES

- [Alves2008] Alves, L. *Estado da Arte dos games no Brasil: trilhando caminhos*. Proceedings of ZON Digital Games 2008, 6-7 Nov, 2008, Porto, Portugal, 9-18.
- [Buschmann96] Buschmann, F. et al., *Pattern-Oriented Software Architecture: A System of Patterns*. John Wiley & Sons Ltd, West Sussex, England, 1996.
- [Cardoso08] Cardoso, B. and Romão, T. *Um Caso Prático de Reabilitação de um Editor de Jogos Móveis Utilizando o Paradigma Model-View-Controller*. Proceedings of Interação 2008, Évora, Portugal, 109-144.
- [Fowler96] Fowler, M. *Analysis Patterns: Reusable Object Models*. Addison-Wesley, Boston, EUA, 1996.
- [Krasner88] Krasner, G. E. and Pope, T., *A cookbook for using the model-view-controller user interface paradigm in SmallTalk*. JOOP, 1(3), Aug, 1988, 26-49.
- [Mell09] Mell, P. Grance, T. 2009. *The NIST Definition of Cloud Computing*. National Institute of Standards and Technology Website.  
<http://src.nist.gov/groups/SNS/cloud-computing>
- [Scott] Scott, J. *Packing Lightmaps*. Website accessed on May 10, 2010.  
<http://www.blackpawn.com/texts/lightmaps/default.html>

# MMORPGs e culturas de convergência

Rui Raposo  
DECA - Universidade de Aveiro  
Campus de Santiago - Aveiro  
raposo@ua.pt

Nídia Salomé Morais  
Escola Superior de Educação de Viseu  
Rua Maximiano Aragão - Viseu  
salome@esev.ipv.pt

## Sumário

*O presente artigo é o resultado de uma breve pesquisa centrada na questão das culturas de convergência no contexto específico dos Massively Multiplayer Online Role Playing Games (MMORPGs). Como conteúdo do mesmo, intenta-se clarificar alguns dos principais conceitos relacionados com a referida área temática, apresentar exemplos que ilustrem, de algum modo, a cultura de convergência nos MMORPGs, bem como se procura identificar alguns factores que poderão influenciar a cultura participativa neste contexto específico.*

## Palavras-chave

*MMORPGs, videojogos, convergência, culturas de convergência, cultura participativa, comunidades online.*

## 1. INTRODUÇÃO

Na imensidão de jogos que se podem encontrar actualmente *online*, um dos géneros de maior sucesso e em constante expansão são os denominados *Massively Multiplayer Online Role Playing Games* (MMORPGs), um tipo de jogo que permite que milhares de utilizadores joguem em simultâneo, em ambientes que poderão ir do simples alfanumérico até cenários virtuais em 3D através da Internet. Cada vez mais, estes jogos caracterizam-se por possuírem altos graus de interactividade, jogabilidade e realismo dos cenários, garantindo desta forma, a sensação de uma maior imersão no ambiente do jogo.

Neste género de jogos, os jogadores interagem entre si numa tentativa de atingir determinados objectivos, sendo que para muitos deles, estes ambientes são como uma segunda vida, onde sociabilizam, partilham, constroem e dão vida a personagens ou a um avatar [Klevjer08]. Koster considera mesmo que um MMORPG é mais do que um simples jogo: *"It's not just a game. It's a service, it's a world, it's a community."* [Koster in Jenkins06:164].

Durante a década de 90, o surgimento deste tipo de jogos para além de introduzir um novo conceito de produção, distribuição e consumo, marca também o início de uma relação mais próxima entre produtores e consumidores desses jogos [Klevjer08], ideia também partilhada por Jenkins: *"Adopting a collaborationist logic, the creators of massively multiplayer online role-playing games [MMORPGs] have already built a more open-ended and collaborative relationship with their consumer base. Game designers acknowledge that their craft has less to do with prestructured stories than with creating the pre-*

*conditions for spontaneous community activities."* [Jenkins06:164]. Com efeito, os designers deste género de jogos têm vindo a reconhecer que a sua responsabilidade enquanto produtores se distancia da mera concepção de histórias com uma estrutura rígida, sendo que nesta lógica de colaboração, têm dado maior importância à concepção de produtos que potenciem a dinamização e a participação em actividades no seio de uma comunidade de jogo *online*.

Para além dos MMORPGs, ambientes especificamente direccionados para uma vertente mais lúdica, têm surgido também mais recentemente outros espaços ou mundos virtuais como, por exemplo, o *Second Life*. Apesar de alguma discussão, em alguns meios, em torno da sua definição, acredita-se que os ambientes como o *Second Life* não devem ser classificados como jogos, na medida em que por exemplo, não estabelecem etapas a ultrapassar, nem objectivos a atingir, tais como ultrapassar níveis ou a obtenção de pontuação. Na opinião de Castronova [Castronova05], estes mundos virtuais devem antes ser entendidos como extensões do mundo real ao simularem algumas das interacções sociais ocorridas nas comunidades reais. Encontram-se, no entanto, semelhanças entre estes mundos virtuais e os MMORPGs no sentido em que partilham uma filosofia idêntica, a de suscitarem a participação activa dos utilizadores em torno de interesses comuns.

Esta possibilidade dos utilizadores poderem dar o seu contributo, de assumirem um papel activo na construção de conteúdos, encontra expressão na ideia de cultura participativa proposta por Jenkins: *"Participatory culture is*

*emerging as the culture absorbs and responds to the explosion of new media technologies that make it possible for average consumers to archive, annotate, appropriate, and recirculate media content in powerful new ways.*" [Jenkins et al.06:8]. Neste sentido, o mesmo autor acredita que as tecnologias *online* têm permitido aproximar os produtores, os *media* e os utilizadores, destacando esta noção de 'utilizadores como produtores' no contexto actual de convergência.

Assim, e mais do que encarar a convergência apenas na sua perspectiva mais tecnológica, isto é, *"the ability of different platforms to carry similar kinds of services or the coming together of consumer devices such as telephone, television and personal computer."* [Schorr in Ludes08:9], importa neste cenário de convergência, atender ao contexto cultural e social em que nos inserimos, ao modo como nos relacionamos com os 'velhos' e com os 'novos *media*'. Assim, numa definição mais abrangente, Jenkins entende o conceito de convergência como sendo *"the flow of content across multiple media platforms, the cooperation between multiple media industries, and the migratory behavior of media audiences who will go almost anywhere in search of the kinds of entertainment experiences they want. Convergence is a word that manages to describe technological, industrial, cultural, and social changes depending on who's speaking and what they thing they are talking about."* [Jenkins06:2-3].

No âmbito do presente artigo, e considerando o contexto específico dos MMORPGs, parece ser pertinente encarar o conceito não tanto na sua perspectiva tecnológica, mas sobretudo sob o ponto de vista em que a convergência *"occurs within the brains of individual consumers and through their social interactions with others."* [Jenkins06:3].

## 2. CULTURAS DE CONVERGÊNCIA COMO CULTURA PARTICIPATIVA NOS MMORPGS

A possibilidade do utilizador participar, no sentido de criar e partilhar novos conteúdos tem ganho novas dimensões decorrentes, entre outros factores, da própria evolução tecnológica e do surgimento de novos serviços, em especial dos serviços que caracterizam esta segunda geração mais social da Web, a ultra-mediatizada Web 2.0<sup>1</sup>. Actualmente, as tecnologias que facilitam a publicação, a partilha, a organização e, sobretudo, a criação de conteúdos para o indivíduo e para a comunidade têm permitido que os utilizadores *"take media into their own hands"*

[Jenkins06:16], fazendo com eles próprios sejam responsáveis por 'moldar' esta cultura dos novos *media* [Jenkins06]. A ideia de que as tecnologias permitem uma alteração do papel do utilizador para que ele próprio seja também produtor de conteúdos, interliga-se directamente com o conceito de cultura de convergência como uma cultura participativa:

*"Convergence culture represents a shift in the way we think about our relations to media, that we are making that shift first through our relations with popular culture, but that skills we acquire through play may have implications for how we learn, work, participate in the political process, and connect with other people around the world."* [Jenkins06:22-23].

Assim, e considerando o contexto de uso de MMORPGs, são evidentes os sinais desta cultura de convergência como uma cultura participativa. Os exemplos que seguidamente serão apresentados neste artigo ilustram que, de facto, é cada vez mais estreita a relação entre produtores e consumidores, pelo que, e tal como sublinha Jenkins, a cultura participativa opõe-se à noção de utilizador passivo e acrescenta que, actualmente, *"rather than talking about media producers and consumers as occupying separate roles, we might now see them as participants who interact with each other."* [Jenkins06:3].

No contexto dos MMORPGs, um dos primeiros exemplos de interacção, participação e trabalho em equipa surgiu em 1993 com o jogo *Doom*. Este jogo podia ser jogado com vários jogadores em redes locais ou através da Internet, sendo que um dos factores cruciais para o crescimento da comunidade *online* em torno deste jogo foi o facto da empresa disponibilizar na Internet os níveis iniciais do jogo para *download* e disponibilizar o código fonte usado na criação de níveis, texturas e sons. Esta estratégia promoveu não só a interacção social, como também a co-produção do jogo entre os utilizadores [Klevjer08].

Em 1997, o jogo *Ultima Online* despoletou o fenómeno das *role-playing communities*, seguindo-lhe exemplos como o *Everquest* [1999], *Anarchy Online* [2001], entre vários outros. A importância destas comunidades nestes contextos é expressa nas palavras de Klevjer:

*"The strength of such communities has significant techno-political implications. As partly independent and self-regulatory centers of consumer power and production of knowledge, they are playing a particular role in ongoing political negotiations over questions of copyright protection, open standards and user control."* [Klevjer08:83].

Considerando a convergência nos *media*, parece pertinente apresentar exemplos de MMORPGs que, de algum modo, são expressão dessa convergência. Assim, como primeiro exemplo refere-se o filme *Matrix*<sup>2</sup> que, para além do filme em si, despoletou também a participação de uma comunidade em torno de um MMORPG. Neste caso, o crescimento dessa comunidade resultou de uma estratégia que envolveu diferentes etapas:

1 "Web 2.0 is the network as platform, spanning all connected devices; Web 2.0 applications are those that make the most of the intrinsic advantages of that platform: delivering software as a continually-updated service that gets better the more people use it, consuming and remixing data from multiple sources, including individual users, while providing their own data and services in a form that allows remixing by others, creating network effects through an "architecture of participation," and going beyond the page metaphor of Web 1.0 to deliver rich user experiences." [O'Reilly, 2005:s.p.].

2 <http://www.mmorpg.com/gamelist.cfm?setview=overview&gameid=88&bhpcp=1>



*“The Wachowski brothers played the transmedia game very well, putting out the original film first to stimulate interest, offering up a few Web comics to sustain the hard-core fan’s hunger for more information, launching the anime in anticipation of the second film, releasing the computer game alongside it to surf the publicity, bringing the whole cycle to a conclusion with the Matrix Revolutions, and then turning the whole mythology over the players of the massively multiplayer online game.” [Jenkins06:97].*

Entre outros, podem ainda referir-se os casos de *Star Wars*<sup>3</sup> e *Harry Potter*<sup>4</sup>. Em ambos os casos é visível a participação dos fãs em fóruns de discussão e em *blogs*, bem como na partilha de vídeos no *Youtube* com dicas para outros utilizadores, com o relato das suas experiências ou mesmo realizando comparações destes MMORPGs com outros jogos do mesmo género.

O envolvimento e a possibilidade dos utilizadores poderem dar o seu contributo tem tido impacto não só no desenvolvimento dos próprios jogos, mas também na criação de novos espaços virtuais, como é exemplo da ‘nova vida’ de *Harry Potter* no *Second Life* ou do mundo virtual da *Star Wars Galaxies*: *“All the settings are referenced from Lucas: the timeline, the planets, our equipment, etc. However, the individual characters and story lines are created by the players. So the filmmaker’s universe is more like a backdrop. Where the stories go and what kind of characters are involved are up to all of us playing in Second Life.”*<sup>5</sup>

Numa tentativa de compreender a cultura participativa nos MMORPGs, os exemplos apresentados permitem ainda identificar um conjunto de factores que poderão influenciar, de algum modo, essa participação. Assim, propõe-se a divisão desses factores em quatro categorias, que seguidamente se explicitam, tentando-se desta forma identificar algumas das razões que poderão promover ou influenciar a cultura participativa no âmbito dos MMORPGs. Trata-se de uma primeira proposta, pelo que um trabalho mais detalhado sobre esta matéria poderia ajudar a afinar alguma da terminologia utilizada, bem como permitiria completar com mais informações os tópicos abordados. Ressalva-se, também, que a ordem com que serão apresentadas essas categorias não significa que tenham maior ou menor importância neste contexto. Deste modo, propõem-se as categorias que se descrevem seguidamente.

## 2.1 Questões normativas

No que se refere às questões normativas, importa mencionar que para que uma comunidade, em determinados contextos, seja criativa e participativa não se deve sentir restringida. Isto é, as questões que regem o mundo real não devem exercer influência na dinâmica da comunidade e do jogo. A este propósito Klevjer refere que *“of partic-*

*ular concern to policy-makers should be the vitality and independence of online gamer communities, which are very distinct practices of collaborative and creative software gaming, creating strongholds of computer literacy and participatory digital culture.” [Klevjer08:83].*

Deverão ser as próprias comunidades a definir o rumo que pretendem dar ao jogo e à sua própria evolução, pelo que as questões normativas não se deverão imiscuir naquilo que é permitido ou não fazer nos jogos e no âmbito das comunidades. Acredita-se que, havendo uma interferência externa, os utilizadores poder-se-ão sentir desmotivados e diminuir a sua participação ou até mesmo abandonar a comunidade.

Estas questões normativas poderiam desencadear, de alguma forma, a discussão em torno da violência e do crime associados a alguns MMORPGs. Convém, no entanto, realçar que os objectivos deste trabalho não passam por essa abordagem, sendo que a perspectiva assumida é sobretudo a de tentar perceber a cultura participativa decorrente da utilização deste género de jogos.

## 2.2 Sentido de pertença

O sentido de pertença a uma comunidade é outro factor que poderá influenciar a participação mais ou menos activa de um membro de uma determinada comunidade de jogo *online*. Tal como no mundo real, muita da participação activa realizada no seio de uma comunidade depende da existência de interesses e objectivos comuns. Especificamente, as ferramentas da Web 2.0 têm permitido que os utilizadores expressem esse sentido de pertença de forma fácil e rápida, seja através de um *post* num blog ou de um vídeo no *Youtube*, entre outros. Tal como afirma Williams [Williams08:28], *“The difference now is that the ability to engage with and combine content across media happens quickly, easily, and is shared by traditional producers and individual members of the audience, with each drawing the interest of fans.”*

Jenkins [Jenkins06] destaca ainda a criatividade e a co-produção que resulta deste sentido de pertença a uma comunidade, onde o poder de decidir a utilização e o significado de um jogo é uma tarefa partilhada pelos fãs e pelos produtores. Neste contexto, podem-se destacar as comunidades criadas em torno de séries de televisão ou filmes, que se transformaram em MMORPGs, como *Star Wars* ou *Harry Potter*, atraindo legiões de fãs para ambientes de interacção que sustentam comunidades *online* devido ao gosto comum partilhado.

## 2.3 Dar e receber [reciprocidade]

Considerando uma comunidade de jogo *online*, acredita-se que o utilizador estará predisposto a contribuir para essa comunidade, desde que dela retire também algum proveito. Esse proveito poderá assumir a forma de novos níveis para o jogo, objectos que permitam melhorar os atributos da sua personagem ou do seu mundo, dicas e códigos para melhorar o seu desempenho, entre outros.

Enquanto alguns autores [Jenkins06] questionam se os jogadores estão realmente dispostos a dedicar mais tempo a criar do que propriamente a jogar, outros defendem que

3 <http://bit.ly/1h4BXv>

4 <http://www.hogwartslive.com>

5 In: <http://bit.ly/d1Bx7D>

*"In games, hard work, patient learning and tedious management are part of the fun."* [Ludes08:11].

Com efeito, existem casos de jogadores completamente altruístas que contribuem de forma bastante intensa sem estarem à espera de receber algo em troca. Para estes jogadores é importante que sintam que *"what they bring to the game makes a difference, not only in terms of their own experiences, but also the experiences of other players."* [Jenkins06:165]. Exemplo disso são as instruções dos jogos elaboradas pelos próprios utilizadores "com descrições fantásticas que ajudam o leitor a desbravar o mundo desse jogo, destacando as estratégias mais eficazes para atingir os seus principais objectivos. Existem centenas de documentos desse tipo *online*, quase todos da autoria de jogadores comuns, que reúnem dicas e técnicas dos amigos e de fóruns de discussão dos jogos. Condensam as ambiguidades e as regras flexíveis desses jogos numa narrativa mais linear, dirigindo-se directamente ao jogador." [Johnson05:157].

No contexto de uma comunidade de jogo *online*, uma das estratégias para a promoção da participação dos seus diferentes membros para 'dar e receber' poderá passar pelo estabelecimento de rácios, onde por cada contribuição existe uma retribuição. Neste caso, utilizadores passivos, ou seja, que não criam para a comunidade poderão ter uma experiência de jogo mais limitada do que o grupo de utilizadores mais activos.

## 2.4 Literacia da informação

A própria evolução das tecnologias *online*, assim como o actual cenário de convergência nos *media* tem potenciado o surgimento novas formas de participação, colaboração e partilha [Alexander08]. Assim, o mesmo autor acredita que as noções de literacia da informação estão cada vez mais ligadas a esta cultura participativa, pelo que, neste contexto, importa promover o desenvolvimento de competências que permitam ao utilizador "identificar, localizar, avaliar, organizar e criar, usar e comunicar a informação de maneira eficaz, para dar resposta a uma questão ou problema." [Wijetunge & Alahakoon05:31].

Os exemplos estudados, no que respeita à cultura participativa em torno dos MMORPGs, ilustram que os jogadores podem desenvolver, muitas vezes em cenários informais, novas aptidões que lhe possibilitam produzir conhecimento e partilhá-lo com os outros, contribuindo desta forma para a inteligência colectiva em ambiente *online* [Jenkins06]. De facto, a literacia da informação de um utilizador poderá ser um dos factores influenciadores da sua participação na comunidade de jogo *online*. Na nossa perspectiva, a cultura participativa em torno dos MMORPGs poderá ser bastante influenciada por aquilo a que Livingstone [Livingstone01] designa como a competência da produção, isto é, a capacidade do utilizador produzir conteúdos para a Internet, bem como interpretar, consumir e desfrutar dos mesmos. No mesmo sentido, Vieira [2008], para além de destacar o acesso e a compreensão, evidencia também a pertinência da criação de conteúdos que "inclui a interacção com os media [...], bem como a produção de conteúdos de media, o que

permite uma melhor compreensão e uma abordagem crítica relativamente a esses mesmos conteúdos." [p.197].

## 3. CONSIDERAÇÕES FINAIS

A pesquisa efectuada no decurso da redacção deste artigo permitiu identificar a existência de fortes indícios de uma cultura participativa no contexto específico dos MMORPGs, pelo que parece ser pertinente continuar a desenvolver investigação nesta área, nomeadamente na identificação das potencialidades da utilização deste tipo de ambientes virtuais noutros contextos como, por exemplo, o de aprendizagem. Deste modo, considera-se que o desafio, neste caso em particular, poderá passar pela construção de narrativas que ajudem os alunos a atingir os objectivos definidos num determinado contexto de aprendizagem, e que garantam, simultaneamente, altos índices de motivação, participação e envolvimento. Embora ainda pendentes de um exercício aprofundado de revisão, avaliação e validação, considera-se que as categorias apontadas poderão servir como pontos de partida para a discussão das estratégias a adoptar para o envolvimento dos utilizadores e da sua transformação em *prosumers* activos.

## 4. REFERÊNCIAS

- Alexander, J. [2008], "Media Convergence: Creating Content, Questioning Relationships", *Computers and Composition*, Vol. 25, pp.1-8.
- Appelgren, E. [2004], "Convergence and Divergence in Media: Different Perspectives", in: *Proceedings of the 8th ICCI International Conference on Electronic Publishing*, Junho 2004, Brasília - DF, Brazil.
- Arnseth, H. [s.d.], "Literacy and media-convergence: Becoming a member of an online gaming community", [consultado em 11 de Novembro de 2008] Disponível na Internet  
<[http://www.pfi.uio.no/konferanse/prof\\_learning/docs/pdf/Paper/Gruppe%204/Arnseth%20-%20paper.pdf](http://www.pfi.uio.no/konferanse/prof_learning/docs/pdf/Paper/Gruppe%204/Arnseth%20-%20paper.pdf)>
- Busà, M. [s.d.], "Will the new media tell a new story?", [consultado em 17 de Novembro de 2008] Disponível na Internet  
<[http://www2.units.it/~didactas/pub/unipd/26Busa\\_Li\\_nguatel.pdf](http://www2.units.it/~didactas/pub/unipd/26Busa_Li_nguatel.pdf)>
- Castronova, E. [2005], "Synthetic Worlds: The Business and Culture of Online Games", [consultado em 12 de Novembro de 2008] Disponível na Internet  
<<http://cs.swan.ac.uk/~csharold/tick/Synthetic.pdf>>
- Chakaveh, S. and Bogen, M. [2007], "Media Convergence, an Introduction", in: *Proceedings of Human-Computer Interaction*, Part III, HCI 2007, pp.811-814.
- Erdal, I. [2007], "Researching Media Convergence and Crossmedia News Production - Mapping the Field", *Nordicom Review*, Vol. 28, pp. 51-61.

- Freitas, S. and Griffiths, M. [2008], "The convergence of gaming practices with other media forms: what potential for learning? A review of the literature." *Learning, Media and Technology*, Vol. 33, pp.11-20.
- Griffiths, M. and Light, B. [2008], "Social networking and digital gaming media convergence: Classification and its consequences for appropriation", *Information Systems Frontiers*, Volume 10, pp.447-459.
- Jenkins, H. [2004], "The cultural logic of media convergence", *International Journal Of Cultural Studies*, Volume 7, pp.33-43.
- Jenkins, H. [2006], *Convergence Culture: where old and new media collide*, New York University Press - New York and London.
- Jenkins, H. et al. [2006], *New Media Literacy White Paper. Confronting the challenges of participatory culture: media education for the 21st century*, MacArthur Foundation.
- Johnson, S. [2005], *Tudo o que é mau faz bem: como os jogos de vídeo, a TV e a Internet nos estão a tornar mais inteligentes*, Lua de Papel: Lisboa.
- Klevjer, R. [2008], "The Cultural Value of Games: Computer Games and Cultural Policy in Europe", in: Ludes, P. [2008], *Convergence and Fragmentation - Media Technology and the Information Society*, Intellect Bristol, UK / Chicago, USA.
- Liestøl, G. [2006], "Dynamics of Convergence & Divergence in Digital Media & Learning", [consultado em 17 de Novembro de 2008] Disponível na Internet <[http://www.editlib.org/toc/index.cfm/files/paper\\_24145.pdf?fuseaction=Reader.DownloadFullText&paper\\_id=24145](http://www.editlib.org/toc/index.cfm/files/paper_24145.pdf?fuseaction=Reader.DownloadFullText&paper_id=24145)>
- Livingstone, S. [2001], "User Empowerment and media competence: Combining protection and education", European Forum on Harmful and Illegal Cyber Content: Self-Regulation, User Protection and Media Competence, [consultado em 20 de Novembro de 2008] Disponível na Internet <<http://www.humanrights.coe.int/media/cyberforum/rep-livingstone.rtf>>
- Ludes, P. [2008], *Convergence and Fragmentation - Media Technology and the Information Society*, Intellect Bristol, UK / Chicago, USA.
- O'Reilly, T. [2005], "[Web 2.0: Compact Definition?](http://radar.oreilly.com/archives/2005/10/web-20-compact-definition.html)", [consultado em 18 de Novembro de 2008] Disponível na Internet <<http://radar.oreilly.com/archives/2005/10/web-20-compact-definition.html>>
- Schwarzwalder, J. [2006], "It's not all games", *Indiana Libraries*, Vol. 25, Number 4.
- Vie, S. [2008], "Digital Divide 2.0: 'Generation M' and Online Social Networking Sites in the Composition Classroom", *Computers and Composition*, Vol. 25, pp.9-23.
- Vieira, N. [2008], "As Literacias e o Uso Responsável da Internet", *Observatorio [OBS\*] Journal*, Vol. 5, pp.193-209.
- Williams, B. [2008], "What *South Park* Character Are You?": Popular Culture, Literacy, and Online Performances of Identity", *Computers and Composition*, Vol. 25, pp.24-39.
- Wijetunge, P. & Alahakoon, U. [2005], "Empowering 8: the Information Literacy model developed in Sri Lanka to underpin changing education paradigms of Sri Lanka", *Sri Lanka Journal of Librarianship and Information Management*, Vol. 1[1], p. 31-41. [consultado em 20 de Novembro de 2008] Disponível na Internet <<http://www.cmb.ac.lk/academic/institutes/nilis/report/InformationLiteracy.pdf>>

Aveiro, 13, 14 e 15 de Outubro de 2010

## **Sessão VII**

### **Acessibilidade**

Aveiro, 13, 14 e 15 de Outubro de 2010



# Caderno Escolar Electrónico Adaptativo

## Resultados Preliminares

Luís Alexandre

LabSI<sup>2</sup>/ESTIG, Instituto Politécnico de Beja, Portugal  
luis.alexandre@ipbeja.pt

Salvador Abreu

Universidade de Évora e CENTRIA FCT/UNL, Portugal  
spa@di.uevora.pt

### Resumo

*Neste artigo descrevemos a construção de uma interface adaptativa para o Caderno Escolar Electrónico, capaz de prever qual a tarefa mais provável que o utilizador irá efectuar no sistema, com base num conhecimento prévio das acções já efectuadas, assim como no tempo e percurso efectuado até determinada altura. Esta adaptabilidade visa a melhoria da interacção entre o sistema e o seu utilizador, reduzindo assim o número de interacções explícitas necessárias para efectuar uma determinada tarefa. Para conseguir prever as tarefas mais prováveis em determinada situação, o sistema recorre a algoritmos de classificação, baseados em árvores de decisão e redes de Markov. Os utilizadores com necessidades especiais, são o principal alvo deste projecto, mas o sistema poderá também dar um grande contributo a outros utilizadores em idade escolar, dado que esta ferramenta favorece a adopção de um processo metódico para recolha e organização de notas em contexto de sala de aula. Por ter sido desenvolvido com recurso a técnicas de desenho para sistemas interactivos, encontra-se totalmente adaptado aos utilizadores alvos.*

### Palavras-chave

*Sistemas Interactivos, Pessoas com Necessidades Especiais, Tecnologias de Apoio, Interacção Pessoa-computador, Inteligência Artificial, Predição de Tarefas.*

## 1. INTRODUÇÃO

O Caderno Escolar electrónico [Alexandre2007] (Figura 1) tem vindo a assumir-se como uma ferramenta de grande utilidade para crianças com dificuldades motoras, dado que para além de permitir a recolha e organização de notas em contexto de aula, armazena e organiza de uma forma lógica todos os documentos referentes a uma determinada disciplina e determinada aula.

O Caderno Escolar electrónico (CE-e) em conjunto com outras tecnologias de apoio, como o Eugénio [Garcia2003] apoiam estas crianças, permitindo que estas possam realizar as tarefas que envolvem escrita, num espaço de tempo muito semelhante ao dispendido pelas crianças sem necessidades especiais. Assim, esta ferramenta ao favorecer a adopção de um processo metódico de recolha de notas em sala de aula poderá introduzir benefícios assinaláveis no sucesso escolar [Boyle2001].

O princípio orientador para o sistema segue uma filosofia DWIM (Do What I Mean) [Teitelman1984], ou seja, o sistema deve comportar-se exactamente de acordo com o objectivo do utilizador, em vez de ter um comportamento coerente com indicações por si dadas, resultantes de um

comando mal executado. Sendo este um objectivo inatingível, é no entanto norteador da nossa intenção.

Mas o CE-e poderá prestar um maior apoio a estes alunos, se reconhecer qual a tarefa em curso, ajudando assim os utilizadores a concluir os passos seguintes da mesma. Em algumas situações é difícil determinar as intenções que se encontram por detrás de uma sequência de acções do utilizador, mas noutros casos, quer recorrendo a conhecimento sobre o domínio [Dix2004] quer observando o comportamento do utilizador em situações anteriores [Maes1994], pode-se conseguir prever qual a operação que pretende realizar. Um caso óbvio será a abertura de uma nova lição no caderno de uma disciplina quando se aproxima a hora dessa aula. Além de prestarem apoio a alunos com dificuldades motoras estes mecanismos também poderão beneficiar alunos com dificuldades cognitivas que com este tipo de ajuda poderão trabalhar de forma mais autónoma em determinadas situações. Este tipo de mecanismos já tem vindo a ser experimentado em algumas ferramentas mais comuns, como por exemplo programas de e-mail [Maes1994] com o intuito libertar o utilizador da execução de tarefas rotineiras através da delegação da responsabilidade pela sua concretização no sistema. Esta abordagem também já foi experimentada no

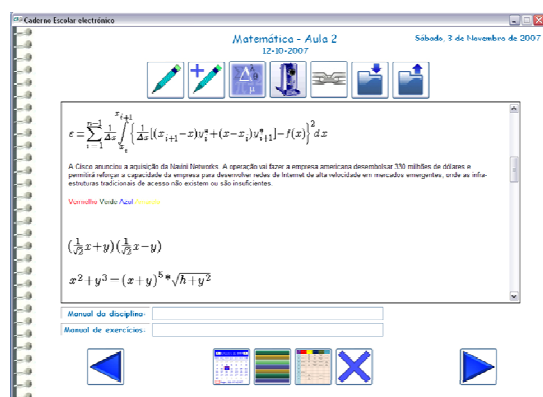


Figura 1 – Interface de notetaking do CE-e

desenvolvimento de um sistema de apoio ao desenvolvimento de configurações de teclado para o Eugénio [Garcia2003, Rodrigues2007].

Este projecto tem como objectivo, implementar e avaliar uma interface inteligente para o CE-e que consiga fornecer um maior nível de suporte aos alunos com necessidades especiais na realização de diversas tarefas.

Mas estas ferramentas necessitam de conter modelos da aplicação e do utilizador [Dix2004] para que seja possível identificar as intenções deste último e assim fornecer ajuda na realização de uma tarefa no sistema.

Assim numa primeira fase deste projecto procedemos à criação de uma base de dados que contem o conhecimento sobre as várias tarefas realizadas pelo utilizador no CE-e, as diferentes acções que as constituem e o contexto onde foram realizadas. Para a caracterização do contexto registamos e armazenamos um conjunto de informação temporal como a data, dia da semana e hora, bem como as acções associadas. Estes dados foram recolhidos no decurso das aulas, com alunos a utilizar o caderno electrónico para o registo das várias actividades lectivas. Os alunos em causa possuem paralisia cerebral, estando apenas afectados na parte física. As actividades, nas quais os registos foram efectuados, decorreram em salas de aula, com turmas entre o terceiro e o sétimo ano de escolaridade, em computadores pessoais com o CE-e instalado. A utilização do sistema foi realizada com recurso aos tradicionais dispositivos de *input/output*, sem a intervenção de quaisquer outros dispositivos adaptados de interacção (e.g. ratos adaptados ou interruptores).

Esta base de dados e de conhecimento foi então utilizada na fase seguinte do projecto, onde estão a ser criados e avaliados mecanismos de apoio ao utilizador na realização de diversas tarefas no CE-e, que entraram em linha de conta com todas as dimensões do contexto. Esta é a principal contribuição deste projecto, dado que se o sistema conseguir interpretar correctamente todas as dimensões do contexto, poderá evoluir de um conjunto de regras *hardcoded* que fornecem um conjunto limitado de adaptações, para um sistema que com a sua contínua utilização conseguirá aprender e criar novas adaptações ao contexto, mais adequadas ao perfil do utilizador. Esta “inteligên-

cia” do sistema só será possível com recurso a algoritmos de aprendizagem automática.

Na secção 2 deste artigo apresentaremos uma revisão do estado da arte, sendo a secção 3 dedicada à descrição da arquitectura do CE-e Adaptativo. Finalmente na secção 4 apresentamos a conclusão e propomos direcções para o trabalho a efectuar no futuro.

## 2. O ESTADO DA ARTE

Hoje em dia vivemos rodeados de sistemas computacionais (e.g. *desktops*, *laptops*, PDAs, entre outros). Vivemos numa época em que todos os aspectos da nossa vida são influenciados ou determinados por decisões ou acções efectuadas por sistemas computacionais.

Este apoio só é possível se o sistema conseguir determinar ou reconhecer o contexto em que tanto a máquina como o seu utilizador estão enquadrados [Blum2005].

Muitos investigadores têm referido a importância e os benefícios que um conhecimento do contexto pode trazer à usabilidade dos sistemas informáticos [Bradley2005], podendo desta forma aumentar e enriquecer a Interação Pessoa-computador.

### 2.1 Context-awareness

A principal ideia por detrás da *context-awareness*<sup>1</sup>, resulta do pressuposto que os computadores podem “sentir” e reagir com base nas alterações do ambiente envolvente. Para que esta reacção seja possível os sistemas computacionais terão de possuir um conjunto de informação proveniente de estímulos e informação anterior que em conjunto com regras pré-estabelecidas, permitirão ao sistema reagir com uma consequência lógica e útil à situação do momento [Schilit1994].

Um sistema *context-aware* pode e deve tentar fazer suposições sobre a envolvente do utilizador. Em [Dey2001] o contexto é definido como “qualquer informação que possa ser utilizada para caracterizar uma qualquer situação de uma entidade”.

Inicialmente o contexto foi apenas entendido como sendo a localização, mas nos últimos anos foi proposto que a localização não fosse entendida como o contexto da situação, mas como um dos componentes do contexto [Dey2001]. Foi ainda proposto que o contexto (adaptação) pudesse ser utilizado para a construção de interfaces adaptativas e inteligentes, nas quais a interacção seja a mais implícita possível.

Muitos dos estudos e tentativas de adaptação ao contexto têm passado pela utilização de dispositivos de localização por satélite. Estes sistemas apenas utilizam um pequeno componente de um relativamente grande universo de possíveis adaptações ao contexto. Um sistema que apenas utiliza a localização para se adaptar ao contexto, não pode ser considerado um sistema *context-aware*, devendo

<sup>1</sup> Conhecimento ou consciência do contexto envolvente.

ser considerado um sistema *location-aware* [Barnard2006].

As aplicações com maior divulgação neste campo são os dispositivos de navegação por GPS utilizados nos veículos automóveis e os dispositivos do tipo guia turístico, normalmente dispositivos do tipo PDA ou *smartphone*. Estes sistemas efectuem uma adaptação dos conteúdos que exibem em função do local onde se encontram.

O *Context-awareness* é maioritariamente utilizado, para além dos sistemas de localização por satélite, em sistemas de computação ubíqua<sup>2</sup>. Neste tipo de sistemas torna-se possível a utilização de sensores e outros dispositivos que captam com precisão o contexto, podendo assim, por exemplo, propor e efectuar alterações de luminosidade ou da abertura das janelas em edifícios, de forma a minimizar os consumos energéticos.

## 2.2 Paradigmas sobre a Adaptabilidade e Previsibilidade

Os primeiros sistemas computacionais tinham como objectivo a resolução de problemas de índole científica ou comercial. Nos dias de hoje todos nós vivemos rodeados de sistemas informáticos, uns mais visíveis, outros mais ubíquos, sendo a interacção cada vez mais intensa, obrigando por isso a cada vez mais acções por parte do utilizador.

Ainda nos primeiros sistemas computacionais, a interface era algo sem grande importância no seu desenvolvimento, sendo esta uma mera ferramenta para fazer a ligação entre a máquina e o utilizador, mas proporcionalmente ao aumento da complexidade dos sistemas, também as interfaces tiveram de evoluir, tornando-se por vezes altamente complexas ou demasiadamente complicadas, ou recheadas de um sem número de *widgets*, que não favorecem a interacção nem favorecem o entendimento das funcionalidades que o sistema e a respectiva interface têm para oferecer.

### 2.2.1 Interfaces Adaptativas e Preditivas

*“Uma interface adaptativa é um sistema computacional interactivo que melhora a sua capacidade de interacção com o utilizador, com base em experiências parciais de interacção com esse mesmo utilizador”* [Langley1997].

Dois dos problemas que mais afectam o desempenho dos utilizadores de sistemas computacionais são a grande quantidade de informação que as interfaces possam apresentar e o facto de se terem de tomar decisões em curtos espaços de tempo [Norcio1989].

Apesar de nos últimos anos terem sido dados passos significativos no que toca à Interacção Pessoa-computador, muitos sistemas computacionais ainda apresentam interfaces de utilizador, complicadas e desajustadas. Esta situação resulta do facto de em muitos destes sistemas a interface não ser tida como parte do sistema, sendo vista

como um mero pacote de *software* do sistema cuja única função é interligar o sistema e o utilizador. Assim, muitas das interfaces não integram no seu desenho princípios de desenho centrado no utilizador, nem estão suficientemente integradas com o sistema, não proporcionando assim uma boa interacção do utilizador com a máquina. Desta desadequação resulta uma não optimização da experiência e da performance do utilizador.

Para ultrapassar este problema foram criadas metodologias de desenho de interfaces, nas quais estas são desenhadas à medida do utilizador ou então o desenho tenta ser o mais adequado possível a todos os eventuais utilizadores do sistema, mesmo que para atingir o grau de adequação óptimo, se tenha de configurar, em utilização alguns aspectos da interface.

Mas o facto de adoptarmos uma metodologia de desenho centrada no utilizador, não nos garante que a interface possa ser igualmente usável por todos os utilizadores, nomeadamente se esses utilizadores forem portadores de algum tipo deficiência. Uma possível resposta para este problema poderá ser a construção de interfaces adaptativas.

A ideia base em que assentam as interfaces adaptativas é muito simples: *“Em vez de ser o utilizador a adaptar-se ao sistema, adapta-se este ao utilizador.”*

Apesar de assentar num pressuposto conhecido, os problemas inerentes são em grande número e complexidade. Uma interface adaptativa tem de se basear em modelos cognitivos e no ambiente circundante [Norcio1989]. Estes modelos tentam explicar o nível de perícia e experiência do utilizador, através de parâmetros do tipo: (i) comandos efectuados, (ii) taxas de erro, (iii) velocidade de digitação, entre outros.

Outro grande problema reside no diálogo, ou seja, o diálogo entre a interface e o utilizador deve ser o mais específico possível para o utilizador em causa. Finalmente, a interface tem obrigatoriamente de ser pensada como um componente do sistema e não como uma mera peça que interliga a camada applicacional e o utilizador.

Uma interface adaptativa pode efectuar as adaptações de duas formas:

- A primeira hipótese de adaptação é uma adaptação relativa ou balizada, ou seja, é efectuada uma adaptação limitada deixando nas mãos dos utilizadores a readaptação da interface, no caso da adaptação efectuada pela mesma não ser satisfatória;
- A segunda hipótese é deixar totalmente nas mãos da interface a responsabilidade pela adaptação.

A segunda hipótese é por muitos considerada como a verdadeira adaptabilidade, sendo no entanto muito difícil de atingir. Uma adaptação totalmente dinâmica apenas conseguirá ser atingida à custa de algoritmos de aprendizagem automática (*machine learning*).

O conceito por trás da aprendizagem automática pode ser por si só considerado como adaptativo, porque os algo-

<sup>2</sup> Modelo de interacção pessoa-computador em que a computação está integrada nos objectos do dia-a-dia.

ritmos deste tipo baseiam-se em informação anterior para tomarem decisões sobre dados actuais.

Contudo uma interface adaptativa também tem os seus problemas e entre outros poderemos referir o facto de o utilizador não conseguir criar um modelo mental do sistema, se a aparência do sistema sofrer muitas e frequentes alterações. Outro potencial problema poderá ser a perda de controlo que o utilizador poderá sentir. Por último e talvez o maior problema das interfaces adaptativas acaba por ser o custo e a complexidade da implementação.

### 2.2.2 Interfaces Atentas e Sugestivas

Outro paradigma sobre a adaptabilidade e previsibilidade, deriva da atenção e da sugestão. Neste caso as interfaces passam a estar atentas e automaticamente sugerem acções aos utilizadores.

Uma interface atenta estabelece automaticamente prioridades e de acordo com as mesmas apresenta ao utilizador a informação ou as opções mais adequadas ao contexto e ao momento de utilização do sistema, para que os recursos, quer do utilizador quer do sistema sejam optimizados, não sobrecarregando assim nenhum dos intervenientes na interacção [Vertegaal2002].

Tal como nas interfaces adaptativas, as interfaces atentas baseiam as suas decisões num conjunto de informações e modelos sobre decisões anteriormente efectuadas pelo utilizador, estados anteriores e estados futuros.

No entanto as interfaces atentas podem-se subdividir em vários subtipos, de uma forma geral esta divisão é efectuada de acordo com a capacidade que as interfaces têm de monitorizar o estado de interacção e atenção do utilizador para com o sistema e a interface.

Nos sistemas cuja interacção é efectuada através de janelas, a atenção do utilizador é captada e comunicada ao sistema de uma forma **explícita** através da organização das janelas, por parte do utilizador de uma forma manual. Por outro lado sistemas que possuam mecanismos sensíveis, tais como sistemas de *eye tracking* para seguirem as acções do utilizador e perceberem onde a atenção do utilizador está centrada, são classificados como sistemas **implícitos**. Uma interface atenta tenderá a utilizar na sua estrutura uma combinação destes dois tipos de interfaces.

Por outro lado temos as interfaces sugestivas, que apenas fornecem ao utilizador pistas sobre a próxima possível acção do utilizador, através do realce de certos componentes da interface [Igarashi2001].

Um utilizador ao interagir com uma interface sugestiva, apenas tem de escolher uma das opções em destaque ou pode ignorar todas as sugestões caso nenhuma corresponda à opção desejada. Por norma as sugestões são geradas por um sistema normalmente apelidado de motor de sugestões, que constantemente observa o estado do sistema, gerando então um conjunto de sugestões que correspondam aos padrões mais próximos do estado actual. Uma interface sugestiva pode ser vista como sendo um sistema de interacção por gestos, ou seja, ao invés de

atualizar o sistema automaticamente, ao ser descoberto um padrão idêntico, a interface apresenta um conjunto de sugestões e pede ao utilizador que escolha uma destas.

Uma interface sugestiva é uma extensão às interfaces preditivas e está a ser encarada por muitos investigadores como a interface do futuro, para a grande maioria dos sistemas, nomeadamente os sistemas WIMP (*Windows, Icons, Menus and Pointing Devices*).

## 3. ARQUITECTURA DO CE-E ADAPTATIVO

A interface do protótipo funcional do CE-e foi desenhada de acordo com uma metodologia de desenho para sistemas interactivos com o focus no utilizador, sendo por isso bastante simples e intuitivo [Alexandre2007]. No entanto como o CE-e pretende fornecer o máximo de ajuda possível em tarefas de *notetaking* e de organização da informação, pretende-se que a interface possa aliviar, o máximo possível, o utilizador de tarefas rotineiras e repetitivas, o que poderá ser particularmente útil para utilizadores com graves problemas motores.

Assim, com base na experiência própria, na opinião de técnicos e de outros peritos, optamos por uma interface do tipo sugestivo, que apresentará ao utilizador a hipótese mais provável, de acordo com os padrões que a base de conhecimento apresente.

### 3.1 A Base de Conhecimento

Tal como foi referido anteriormente, para sistema poder prever ou sugerir a acção mais provável para um determinado contexto, tem de se socorrer de conhecimento passado e modelos de interacção. Para a construção desta base de conhecimento, recorre-se frequentemente a sistemas que registam as acções do utilizador – *logging*.

Há já algum tempo que investigadores que trabalham em sistemas de Comunicação Aumentativa e Alternativa (CAA), tentam medir o impacto dos avanços tecnológicos no acto de comunicar com recurso a este tipo de sistemas. Assim, de forma a facilitar a análise dos dados recolhidos, com base em sistemas de *logging*, foi definido um formato padrão que permite efectuar a análise aos dados de uma forma sistemática, o *Universal Logging Format* [Lesh2000].

A estrutura do ficheiro de logs proposta em [Lesh2000], apresenta três partes distintas: (i) Cabeçalho que especifica o conteúdo e o formato dos registos; (ii) Corpo que é composto por  $n$  linhas, cada uma representativa dos dados relativos a um *log*; (iii) Secção de análise, sendo esta opcional, onde se podem colocar algumas estatísticas dos *logs*.

cid	name	type	notnull	def_value	pk
0	ID	INTEGER	0		1
1	date	NVARCHAR(40)	0		0
2	weekDay	NVARCHAR(20)	0		0
3	time	NVARCHAR(50)	0		0
4	form	NVARCHAR(50)	0		0
5	discipline	NVARCHAR(50)	0		0

Figura 2 – Estrutura dos logs do CE-e





```

1 @relation CE-e_laura_2010-03-24
2
3 @attribute Day (Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday)
4 @attribute Hour numeric
5 @attribute Form (InsertEditNote, Form1, FormAgenda, booksForm, BooksForm, ScheduleViewForm, NotesForm)
6 @attribute Discipline ("Apoio ao Estudo ", "Inglês ", "Estudo do Meio ", "Matemática ", "Língua Portuguesa ", "")
7 @attribute Event ("formatar texto alinhamento centro", "formatar texto cor", "formatar texto highlight", "formatar texto tamanho fonte",
8 "formatar texto sublinhado", "formatar texto bold", "novo evento", "navegar vista configuracoes", "editar nota texto",
9 "cancelar escolha nota texto", "configurar manual disciplina", "apagar nota texto", "editar hiperligacao", "escolha nota texto",
10 "formatar texto italico", "abrir manual disciplina", "download ficheiro", "inserir hiperligacao", "iniciar nova aula",
11 "abrir manual exercicios", "terminar aplicacao", "iniciar aplicacao", "navegar vista livros", "abrir caderno",
12 "escolha impressao objecto", "cancelar escolha impressao objecto", "impressao nota-link-img", "navegar aula anterior",
13 "navegar form agenda", "navegar aula posterior", "inserir nota texto")
14
15 @data
16 "Thursday",100331,"FormAgenda","", "navegar vista livros"
17 "Thursday",100334,"BooksForm","", "abrir caderno"
18 "Thursday",100347,"NotesForm","Estudo do Meio ", "navegar vista livros"
19 "Thursday",100350,"BooksForm","", "abrir caderno"
20 "Thursday",100509,"NotesForm","Língua Portuguesa ", "iniciar nova aula"
21 "Thursday",100639,"NotesForm","Língua Portuguesa ", "inserir hiperligacao"
22 "Thursday",100852,"FormAgenda","", "navegar vista livros"
23 "Thursday",100855,"BooksForm","", "abrir caderno"
24 "Thursday",100910,"FormAgenda","", "navegar vista livros"
25 "Thursday",100912,"BooksForm","", "abrir caderno"

```

Figura 5 – Aspecto parcial do ficheiro submetido à ferramenta Weka

base de conhecimento e apresentando de seguida a hipótese que mais se adequa ao contexto em causa e à componente temporal do momento.

Os primeiros testes efectuados, com recurso a uma árvore de decisão C4.5 [Quilan1993] (Figuras 4 e 5), demonstraram que poderiam ser atingidos resultados com uma percentagem mínima de acertos na ordem dos 50%, valor que justifica que se desenvolva um esforço de implementação, dado que o sistema possui, neste momento 40 classes. No entanto destas 40 opções que o sistema possui, dependendo do ecrã onde o utilizador se encontra, apenas são apresentadas entre 3 a 17 opções, o que em termos médios nos dará a cada momento 7 opções possíveis, sendo que 50% de 7 dará aproximadamente 4. Uma percentagem de acertos de 4 em 7 é um resultado bastante encorajador.

Outros algoritmos de classificação foram testados em busca de melhores resultados ou de resultados que confirmassem os obtidos com recurso a árvores de decisão, nomeadamente os Modelos "Escondidos" de Markov (do Inglês *Hidden Markov Models*) [Rabiner1989]. À semelhança das árvores de decisão, os modelos "escondidos" de Markov, são particularmente adequados para conjuntos discretos de dados. Os resultados obtidos com este algoritmo foram igualmente encorajadores, encontrando-se ao nível dos obtidos pela árvore de decisão C4.5. Ao contrário das tradicionais redes de Markov, em que o estado seguinte depende apenas do estado actual do sistema, os modelos "escondidos" de Markov tentam inferir sobre qual a sequência mais lógica de acontecimentos ou estados que produziram um determinado estado ou objectivo, dado que em muitas situações os sistemas não conseguem determinar certos estados - estados "escondidos". Este algoritmo acabou por ser o escolhido, devido à sua mais rápida implementação no sistema, através de classes no código existente do CE-e.

O motor de inferência do sistema lidará com um conjunto pequeno de atributos, como se pode ver na figura 4 o motor terá apenas quatro atributos que necessita de analisar, dado que o quinto atributo é o atributo classe, que servirá para treinar o classificador. No entanto este classificador terá de lidar com um sistema multi-classe, sendo

as classes em grande número. A matriz de confusão construída pela ferramenta Weka (ver figura 3) demonstra de uma forma bastante clara o número e dimensão das classes, sendo este um problema para o classificador, dado que a árvore resultante possuirá um grande número de nós e folhas, aumentando assim a probabilidade de erro na sugestão da hipótese mais provável. No entanto os 52% de acertos conseguidos garantem que pelo menos uma em cada duas sugestões estará correcta reduzindo assim a carga sobre o utilizador, na realização de tarefas que acrescentam valor ao trabalho do utilizador no sistema.

### 3.3 A codificação

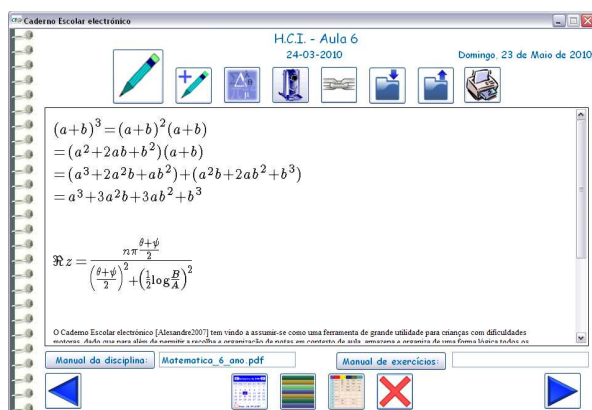
A codificação do motor de inferência desenrola-se em torno de dois algoritmos, uma árvore C5, que deriva da árvore C4.5 de [Quilan1993] e, um algoritmo baseado em modelos "escondidos" de Markov. Apesar de o primeiro ter apresentado melhores resultados nos testes realizados na ferramenta *Weka*, a codificação do sistema avançou com recurso ao segundo algoritmo. Esta decisão foi motivada pela urgência de possuímos um protótipo pronto antes do início do ano escolar, para que se pudesse comprovar a real utilidade da Interface Adaptativa do Caderno Escolar electrónico.

Apesar do motor de inferência deste primeiro protótipo da Interface Adaptativa do Caderno Escolar electrónico, ter sido implementado com recurso ao algoritmo baseado em modelos "escondidos" de Markov, temos como objectivo final a implementação do sistema com recurso ao algoritmo baseado na árvore de decisão C4.5/C5, dado que esta é mais rápida e ligeiramente mais eficaz na procura de padrões e na classificação em classes.

Ao nível da codificação, estas implementações seguem metodologias diferentes, enquanto que a árvore de decisão será implementada através da concepção de uma DLL baseada em código disponibilizado em *open source* por Quilan [Quilan1993], a segunda seguiu uma implementação por classes integradas no código fonte do CE-e.

Assim, foi possível implementar e testar o sistema em busca de erros ou problemas de implementação, em tem-





**Figura 6 – Protótipo do sistema com o motor de inferência a sugerir a opção mais provável**

po útil, antes do início do ano escolar no ensino básico, onde o sistema será efectivamente colocado à prova.

Após a prova de conceito ter sido realizada, a DLL baseada na árvore de decisão de [Quilan1993], continuará a ser desenvolvida. Este esforço justifica-se pelo facto de uma DLL ser independente do sistema e portátil, permitindo assim a sua reutilização noutros sistemas, para além de ser mais fácil a construção de novas versões do CE-e, com e sem interface inteligente, sem que esta situação origine a construção de duas versões, incompatíveis.

Com a introdução do motor de inferência no protótipo funcional do CE-e, o sistema teve de sofrer algumas alterações, não visíveis, que apenas se reflectiram no seu código fonte, para que os vários componentes da sua interface possam ser alterados dinamicamente por indicação do motor de inferência da interface.

#### 4. CONCLUSÃO E TRABALHO FUTURO

Neste artigo propusemos e apresentamos uma interface adaptativa para uma ferramenta que pretende ser uma alternativa aos tradicionais cadernos escolares, para estudantes com necessidades especiais, em particular aqueles que apenas possuem problemas físicos, que lhes restringe a mobilidade e a habilidade para manipular os tradicionais mecanismos de *input*.

Actualmente dispomos de um protótipo do CE-e com a interface adaptativa, em fase final de desenvolvimento. Este protótipo inicial apenas será utilizado para testes e prova de conceito, que estão a decorrer.

A interface faz sugestões ao utilizador com recurso à análise de uma base de conhecimento, que contem informação sobre acções anteriormente efectuadas no sistema. A análise é efectuada com base em técnicas de aprendizagem automática (problema de classificação multi-classe), através de algoritmos baseados em *Redes de Markov*, numa primeira fase e em árvores de decisão na altura da sua implementação final.

Com vista à construção de uma base de conhecimento que conseguisse responder às questões colocadas durante a fase de análise e decisão, o mecanismo de registo de *logs* passou por uma base de dados relacional, com recurso à

tecnologia SQLite, dado que a interrogação de bases de dados em SQL é uma prática bastante eficiente. De igual forma o registo de logs num sistema deste é bastante simples e eficiente, não apresentando qualquer tipo de sobrecarga, quer para o caderno electrónico, como para o sistema operativo.

Como se pretende que o sistema vá aprendendo ao longo do tempo, o ficheiro de *logs* poderá crescer indefinidamente tornando a sua análise pelo motor de inferência da interface do CE-e, muito lenta. Para ultrapassar este potencial problema, o ficheiro *logs* poderá ver o seu crescimento limitado. Esta limitação no crescimento, poderá ser regulada com recurso ao tempo, sendo os *logs* com datas posteriores a trinta dias (efectivamente registados no mesmo) eliminados. Se verificarmos que este período é demasiado curto, ou longo demais, esta regra será reajustada.

No entanto, a hipótese de existir uma limitação temporal dos *logs*, pode ainda sofrer alterações dado que estamos a analisar qual a opção que trará maiores probabilidades de acertos. Se por um lado o sistema apenas conhecer os últimos trinta dias de utilização, terá um conhecimento temporalmente limitado, mas esse conhecimento poderá incidir sobre as opções que o utilizador tem utilizado com maior frequência. Por outro lado e, partindo da hipótese que os *logs* podem crescer indefinidamente, o sistema terá uma visão global de todas as acções e escolhas efectuadas pelo utilizador, o que poderá trazer maiores probabilidades de acerto. Esta decisão só poderá ser tomada, após um período de utilização e testes de duração não inferior a seis meses, que demonstrará qual o melhor tipo de base de conhecimento, “com esquecimento” ou “sem esquecimento”

Uma eventual solução para a recolha contínua de logs, poderá passar pelo seu processamento em *offline*, o que levará à construção de um índice de acções, à semelhança do que se passa num sistema de recuperação de informação em bases de texto [Manning2008].

Outro objectivo, inerente a esta proposta, é dotar o sistema de um conjunto de regras fixas, que em caso de não correspondência entre a mesma e o resultado obtido pelo motor de inferência, se sobreporá à inferência do mesmo. Como exemplo deste tipo de regras temos o horário escolar que poderá fornecer informações sobre qual a tarefa mais provável a executar no sistema, de acordo com a componente tempo fornecida pelo sistema operativo.

O mecanismo de sugestão coloca em realce a opção mais provável. Este realce está aliado a uma tecla de atalho, o que evita um movimento de rato para selecção da opção.

O CE-e pretende ser uma alternativa aos tradicionais cadernos escolares. Se a introdução de uma interface adaptativa provar ser uma mais-valia, os estudantes com dificuldades físico-motoras, terão ao seu dispor uma ferramenta que os ajudará ainda mais na organização de todas as tarefas escolares que envolvam escrita e na organização de todos os materiais escolares.

Nesta altura o CE-e possui uma interface adaptativa que consegue prever qual a acção mais provável para um determinado contexto, o que veio reforçar o principal objectivo do CE-e: Ajudar e aliviar os estudantes com necessidades especiais de tarefas rotineiras, que não acrescentam mais-valias ao trabalho que têm de efectuar, de forma a ajudar na sua integração no ensino regular.

Por último gostaríamos de agradecer a preciosa colaboração do corpo docente do Instituto Politécnico de Beja afecto ao Laboratório de Sistemas de Informação e Interactividade (LabSI<sup>2</sup>), pelos contributos e ideias dadas para este projecto.

## 5. REFERÊNCIAS

- [Alexandre2007] Alexandre, L., Garcia, L., Bruno, L.. Development of an electronic scholar notebook for students with special needs. DSAI2007, Vila Real, Portugal. 2007.
- [Alpaydin2004] Alpaydin, E. Introduction to Machine Learning. MIT Press. 2004.
- [Barnard2006] Barnard, L., Yi, J., Jacko, J., Sears, A. Capturing the effects of context on human performance in mobile computing systems. 2006.
- [Blum2005] Blum, M.L. Real-time context recognition. Master's thesis, Department of Information Technology and Electrical Engineering, Swiss Federal Institute of Technology Zurich (ETH), 2005.
- [Boyle2001] Boyle, J., Weishaar, M. The effects of strategic notetaking on the recall and comprehension of lecture information for high school students with learning disabilities. Learning Disabilities Research & Practice, 16(3), pages 133–141, 2001.
- [Bradley2005] Bradley, N., Dunlop, M. Toward a multidisciplinary model of context to support context-aware computing. Human Compute-Interaction, 20:403 – 446, 2005.
- [Dey2001] Dey, A. Understanding and using context. Personal Ubiquitous Computing, 51:4 – 7, 2001.
- [Dix2004] Dix, A., Finlay, J., Abowd, G., Beale, R. Human Computer Interaction, 3rd Edition. Prentice Hall, 2004.
- [Garcia2003] Garcia, L. Conceção, implementação e teste de um sistema de apoio à comunicação aumentativa e alternativa para o português europeu. Tese de Mestrado. Universidade Técnica de Lisboa, Instituto Superior Técnico, 2003.
- [Garner1995] Garner, S.R. Weka: The Waikato environment for knowledge analysis. In Proc New Zealand Computer Science Research Students Conference, pages 57-64, University of Waikato, Hamilton, New Zealand, 1995.
- [Gonçalves2003] Gonçalves, M., Panchanathan, G., Ravindranathan, U., Krowne, A., Fox, E. JCDL'03 Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital libraries. 2003.
- [Igarashi2001] Igarashi, T., Hughes, J. A suggestive interface for 3D drawing. Proceedings of the 14th annual ACM symposium on User interface software and technology. Pages: 173 – 181. Orlando, Florida. 2001.
- [Langley1997] Langley, P. Machine Learning for Adaptive User Interfaces. Proceedings of the 21st Annual German Conference on Artificial Intelligence: Advances in Artificial Intelligence. Pages: 53 – 62. 1997.
- [Lesh2000] Lesh, G., Moulton, B., Rinkus, G., Higginbotham, G. A Universal Logging Format for Augmentative Communication. Enkidu Research, Inc. Lockport, NY. 2000.
- [Maes1994] Maes, P. Agents that reduce work and information overload. Communications of the ACM 37(7), 1994.
- [Manning2008] Manning, C.D., Raghavan, P., Schütze, H. An Introduction to Information Retrieval. Cambridge University Press. 2008.
- [Martinez2006] Frias-Martinez, E., Chen, S., Liu, X. Survey of Data Mining Approaches to User Modeling for Adaptive Hypermedia. IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews, Volume 36, No. 6. 2006.
- [Norcio1989] Norcio, A., Stanley, J. Adaptive human-computer interfaces: A literature survey and perspective. IEEE Transactions on Systems, Man and Cybernetics, 19, No.2:399 – 408, 1989.
- [Quilan1993] Quinlan, R. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA. 1993.
- [Rabiner1989] Rabiner, Lawrence R. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE 77. 1989.
- [Rodrigues2007] Rodrigues, N. Desenvolvimento de mecanismo de interacção preditiva para aumentar o desempenho de tarefas. Projecto de Licenciatura. Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Beja, 2007.
- [Schilit1994] Schilit, B., Adams, N., Want, R. Context-aware computing applications. IEEE Workshop on Mobile Computing Systems and Applications (WMCSA'94), 1:89 – 101, 1994.
- [Schilit1994-2] Schilit, B., Theimer, M. Disseminating active map information to mobile hosts. IEEE Network, pages 22 – 32, 1994.
- [Teitelman1984] Teitelman, W. A tour through cedar. Proceedings of the 7th international conference on Software engineering. Páginas: 181 – 195. Orlando, Florida, United States. 1984.
- [Vertegaal2002] Vertegaal, R. Designing Attentive Interfaces. ETRA'02. New Orleans - Louisiana, USA - 2002

# Promovendo a Colaboração entre Crianças do Pré-Escolar através da Realidade Aumentada

Sofia Pessanha  
CCM - Universidade da Madeira  
9000-390 Funchal  
sofia.pessanha@gmail.com

Pedro Campos  
Univ. da Madeira e INESC ID Lisboa  
Funchal / Lisboa  
pedro.campos@inesc-id.pt

## Sumário

*A importância da familiarização das crianças com a tecnologia é cada vez mais notável: quer porque esta faz parte inquestionável do mundo que a rodeia, quer pela relevância educativa das experiências que lhe pode proporcionar.*

*As crianças do pré-escolar são um grupo de utilizadores especiais, uma vez que se encontram numa fase inicial das suas vidas, onde têm de aprender a viver em sociedade, isto é, aprender a ouvir e respeitar as opiniões dos outros, partilhar os mesmos objectos e também a ajudar-se mutuamente. Neste estudo, foi desenhado e desenvolvido um jogo em Realidade Aumentada para o ensino de conceitos simples que promove acções colaborativas entre as crianças. Todo o design do jogo foi devidamente acompanhado por educadoras.*

*Os resultados sugerem que o jogo é eficaz para obter níveis altos de concentração, envolvimento, motivação e curiosidade que levam a que haja uma maior colaboração entre as crianças, particularmente quando o feedback (resultado) do jogo é fornecido de forma imediata.*

*Embora já existam diversas soluções voltadas para o ensino, tais como jogos sobre a matemática, geometria e sobre o sistema solar, existe claramente uma falta de soluções e estudos sobre a aplicação desta tecnologia com crianças do pré-escolar (3 aos 5/6 anos). Esta foi uma das principais razões que nos levou a desenvolver esta investigação.*

## Palavras-chave

*Realidade aumentada; Sistemas de aprendizagem interactivos; Aprendizagem colaborativa; Questões pedagógicas; Interfaces Tangíveis; Interfaces Multimodais.*

## 1. INTRODUÇÃO

Através das novas tecnologias podem ser desenvolvidos jogos educativos para as crianças, uma vez que estas estão cada vez mais em contacto com as tecnologias existentes em seu redor. O poder dos jogos digitais como uma ferramenta educativa tem sido amplamente reconhecido [Mayo07] e estes podem ser utilizados para ensinar diversos temas a alunos de várias idades [Gibson03].

A título de exemplo, [Gibson03] descreve um jogo destinado a ensinar diferentes temas a crianças de diversas idades. Belotti e colegas [Belotti09] descrevem um jogo educativo utilizando uma abordagem estado da arte em termos de desenvolvimento comercial, enriquecendo o ambiente com instâncias dos módulos educativos desenvolvidos. Os objectivos didácticos destas abordagens são essencialmente explorar o potencial dos computadores para atingir uma faixa demográfica tradicionalmente resistente ao processo de aprendizagem.

Numa linha de investigação mais específica, têm sido desenvolvidos jogos utilizando técnicas de Realidade Aumentada (RA) no contexto da sala de aula. Desde matemática de Liceu até geometria [Kaufmann02], sis-

temas solares interactivos [Medicherla10] entre outros, o âmbito de aplicação tem sido relativamente amplo.

Contudo, existe uma clara lacuna no que diz respeito a soluções e estudos destas tecnologias com crianças do pré-escolar, cujas idades variam entre os três e os cinco anos. Estas crianças têm capacidades e objectivos de aprendizagem bastante diferentes das outras.

Os objectivos deste estudo, consistem essencialmente em explorar o potencial das novas tecnologias direccionadas às crianças do pré-escolar. Para tal, ao invés de ser desenvolvido um programa de computador que utiliza técnicas de introdução de dados convencionais (rato e teclado), este trabalho apresenta uma nova interface para a aprendizagem de conteúdos educativos especificamente concebida para crianças do pré-escolar (na faixa etária dos três aos cinco anos de idade).

A ideia consiste essencialmente em adicionar elementos virtuais a algo do mundo real, realizando esta interacção com as nossas próprias mãos. Assim, para utilizar este sistema, não é necessário que as crianças tenham experiência com computadores, nomeadamente facilidade de utilização do rato e do teclado, que é uma habilidade que

nem todas as crianças possuem nestas idades. A interface, consiste, essencialmente numa simbiose dos tradicionais jogos de tabuleiro com as novas tecnologias.

O racional da nossa abordagem subdivide-se em dois aspectos. Primeiro, Papert refere que “a aprendizagem é mais eficaz quando o aprendiz se envolve voluntariamente no processo” [Papert96]. Motivar os aprendizes é portanto um factor crucial para aumentar a possibilidade de acção e descoberta, o que por sua vez aumenta a capacidade daquilo que alguns investigadores chamam a “capacidade de aprender a aprender”. A RA [Shelton02] enquadra-se bem neste paradigma como técnica de suporte, já que promove níveis altos de motivação e envolvimento.

Em segundo lugar, as crianças do pré-escolar têm objectivos e necessidades educativas que são muito diferentes das restantes, muito mais baseadas em noções colaborativas, como o respeito pelos colegas, interacção com os outros, resolução colaborativa de problemas outros valores da vida em sociedade. O design de novos jogos para o pré-escolar deve, por estes motivos, promover bons níveis de colaboração de forma eficaz.

Neste artigo, é apresentada uma interface tangível que vai ajudar a perceber de que forma podemos obter um maior número de acções colaborativas entre as crianças, através de duas versões de um jogo educativo.

## 2. TRABALHO RELACIONADO

A tecnologia actual fornece novas possibilidades emocionantes para aproximar as crianças aos conteúdos digitais. Existem numerosas áreas onde a RA pode ser aplicada, áreas que vão desde o sério até ao entretenimento. Assim, o processo de visualização e manipulação de objectos virtuais em ambientes reais pode ser encontrado em inúmeras aplicações, sobretudo em educação e formação, onde o potencial é grande, já que é necessário ter recursos que permitam uma melhor visão do objecto de estudo. Outras aplicações incluem a criação de ambientes colaborativos em RA, onde cada utilizador visualiza e interage com elementos reais e virtuais.

Dado o âmbito do nosso trabalho, dividimos a revisão da literatura em dois aspectos: a utilização de tecnologia RA na sala de aula; e abordagens focadas na promoção da colaboração na sala de aula através de novas tecnologias, não necessariamente baseadas em RA.

### 2.1 Realidade Aumentada na Sala de Aula

A utilização de sistemas de RA em contextos educativos não é nova, *per se*. Shleton e Hedley descrevem um projecto de investigação onde utilizaram RA para ajudar a ensinar estudantes de licenciatura em Geografia conceitos acerca das relações Sol-Terra [Shelton02]. Examinaram mais de trinta alunos que participaram num exercício de RA contendo modelos destinados a ensinar conceitos de rotação, revolução, solstício/equinócio e variação de luz e temperatura em função das estações do ano, tendo encontrado uma melhoria significativa na compreensão dos estudantes após o exercício de RA, assim como uma redução do número de conceitos mal aprendidos.

Outras importantes conclusões acerca deste sistema foram o facto de as interfaces RA não se limitarem a alterar o mecanismo de distribuição do conteúdo educativo: elas podem fundamentalmente alterar a forma como o conteúdo é compreendido através de uma combinação única de informação visual e sensorial que resulta numa poderosa e cognitiva experiência de aprendizagem [Shelton02].

As simulações em ambientes virtuais têm vindo a transformar-se numa importante ferramenta de investigação para os educadores [Tettegah06]. Em particular, a RA tem sido utilizada para ensinar modelos físico-químicos [Tettegah06]. Schrier avaliou as percepções relativas a duas representações na aprendizagem dos aminoácidos [Schrier06]. Os resultados mostraram que alguns estudantes gostaram de manipular os modelos RA através de rotações nos marcadores para observar as diferentes orientações dos objectos virtuais.

O vasto leque de aplicações educativas RA também se estende à Física pura. Duarte et al. [Duarte05] usam RA para apresentar dinamicamente informação associada à mudança de cenário utilizado no mundo real. Neste caso, os autores realizaram uma experiência no campo da física para mostrar informação que varia no tempo, como a velocidade e a aceleração, a qual pode ser estimada e mostrada em tempo real. A visualização de dados reais e estimados durante a experiência, assim como o uso de técnicas de RA, provou ser bastante eficiente, já que as experiências puderam ser detalhadas e interessantes, promovendo os mecanismos cognitivos da aprendizagem.

Min-Chai Hsieh and Jiann-Shu Lee [Hsieh08] desenvolveram um sistema chamado ARELS (*Augmented Reality English Learning System*) destinado a ajudar a aprendizagem de inglês das crianças do pré-escolar. Criaram um marcador para cada letra do alfabeto e associaram cada um desses marcadores a um modelo 3D de uma palavra iniciada por essa letra. O modelo RA mostrava também texto, imagens, músicas, vídeos e animações. Também foi proposto um método para reduzir a complexidade e aumentar a capacidade de desenhar marcadores RA [Hsieh08].

A utilização da RA na educação formal também pode constituir um componente chave nos ambientes de aprendizagem do futuro. Estes ambientes serão abundantemente povoados com aplicações misturando avanços em hardware e software.

### 2.2 Promovendo a Colaboração

A promoção de comportamentos colaborativos é crucial no contexto educativo do pré-escolar. Por isso, analisamos de forma breve as abordagens que utilizam a tecnologia como forma de alcançar níveis mais altos de colaboração dentro da sala de aula.

É sabido que as crianças comunicam e aprendem através de brincadeiras e exploração [Sutton86]. Através da interacção social e através da imitação, as crianças adquirem novas capacidades e aprendem a colaborar. Isto também é verdade quando as crianças trabalham com computadores. Utilizando interfaces tradicionais, baseadas em

teclado e rato, e mesmo tendo em consideração que duas ou mais crianças podem colaborar verbalmente, apenas uma criança está em controlo do computador ao mesmo tempo. O reconhecimento de que o trabalho de grupo sobre um único computador é desejável levou ao desenvolvimento de software e hardware especificamente concebidos para dar suporte a interacção por vários utilizadores ao mesmo tempo. Esta interacção conduziu a significativos aumentos dos níveis de aprendizagem [Inkpen97].

Stewart et al. observaram que as crianças com acesso a múltiplos dispositivos de entrada apreciavam uma experiência de utilização melhor, tendo estes investigadores notado aumentos nas incidências de interacções estudante-estudante e estudante-professor, assim como um carácter mais colaborativo nas interacções. As crianças também pareciam ter apreciado mais a experiência comparativamente a observações anteriores em sistemas standard [Stewart98].

Também existem estudos sobre o design de interfaces com o utilizador dedicadas à colaboração entre crianças [Africano04]. Alguns resultados apresentam sistemas que efectivamente suportaram colaboração e interactividade que foi apreciada pelas crianças, as quais se sentiram imersas no jogo [Africano04].

Kannetis e Potamianos investigaram a forma como a fantasia, curiosidade e desafio contribuem para a experiência de utilização em jogos multimodais para crianças do pré-escolar, o que é particularmente relevante para a nossa investigação. Descobriram que a fantasia e a curiosidade se encontram correlacionadas com os níveis de entretenimento das crianças, ao passo que os níveis de dificuldade parecem depender das capacidades e preferências individuais de cada criança [Kannetis09].

Um aspecto que foi tido em conta no design do nosso jogo de RA para pré-escolar foi o facto de as crianças se envolverem mais quando as interfaces são multimodais e contêm elementos que despoletam a sua curiosidade.

### 3. O JOGO

O sistema desenvolvido é baseado num tabuleiro de madeira contendo nove divisões onde as crianças deverão colocar as peças do jogo que consistem em marcadores de Realidade Aumentada (RA).

As educadoras de infância tiveram um papel activo no design do sistema e forneceram um objectivo de aprendizagem que consistiu no estudo dos seres vivos e dos meios onde vivem (mar, ribeira, terra e ar). Cada divisão do tabuleiro contém uma imagem correspondente a um ou mais meios. Cada peça do jogo corresponde a um animal em 3D que pode ser livremente manipulado. As peças foram construídas com um material esponjoso, próprio para ser utilizado por estas crianças. Este tabuleiro contém uma câmara numa posição fixa, responsável por processar a informação vídeo em tempo real. A Figura 1 mostra o sistema, que pode ser conectado a qualquer computador.

O objectivo do jogo consiste em conseguir colocar todas as peças (marcadores) nos respectivos compartimentos correctos do tabuleiro. O jogo termina quando o tabuleiro estiver preenchido de uma forma correcta (existe mais de uma possibilidade).



**Figura 1: O sistema com recurso a um ecrã LCD.**

O feedback sobre o resultado do jogo, é fornecido quando um marcador especial, concebido apenas para esse propósito, é colocado no respectivo compartimento de nome “Resposta”.

Duas versões diferentes do jogo foram desenvolvidas para estudar o impacto do feedback nos níveis de colaboração entre as crianças: uma versão onde o feedback é fornecido a qualquer altura do jogo, ou seja, sempre que uma criança coloca o marcador especial para saber a resposta; e uma versão onde o feedback é apenas dado no final do jogo, isto é, quando todas as peças (seres vivos) tiverem sido colocadas no tabuleiro.



**Figura 2: Projecção do jogo, mostrando o feedback.**

O feedback é mostrado da seguinte forma: os seres vivos que estão colocados correctamente têm um contorno verde (árvore, lobo-marinho, estrela-do-mar e gaivota), e os que estão incorrectos têm um contorno vermelho (golfinho, borboleta, peixe e sapo), de acordo com a Figura 2. De acordo com as sugestões das educadoras, decidimos acrescentar feedback sonoro. Esse feedback consiste em



frases pré-gravadas, tais como: “Não está certo, tenta outra vez!”, “Muito bem, conseguiu!” (entre outras). Assim encoraja-se as crianças, principalmente quando o reforço positivo é dado com aplausos.

#### 4. AVALIAÇÃO

Uma vez que o impacto de uma tecnologia não pode ser plenamente compreendido sem considerar o contexto educativo como um todo [Salomon90], realizamos um método misto para este estudo. Uma combinação de métodos quantitativos (pré e pós-testes realizados em dias diferentes) e qualitativos (observações, questionários e gravações vídeo) que permitiram uma compreensão fundamentada do impacto do jogo na (i) aprendizagem, (ii) motivação e nos (iii) níveis de colaboração. Apesar de apresentarmos os resultados para os três itens (efectividade da aprendizagem, os níveis de motivação e os níveis de colaboração), o nosso principal foco foi sobre os níveis de colaboração, ou seja, queremos investigar de que forma um jogo como este pode promover níveis mais elevados de comportamentos de colaboração.

Os testes foram realizados em diferentes escolas e num museu, com um total de 101 crianças.

As crianças foram agrupadas em grupos de quatro ou cinco (se necessário). Foi ainda utilizado um projector para permitir a participação de toda a turma na actividade. A Figura 3 ilustra, a título de exemplo, um dos testes realizados.

#### 5. RESULTADOS

Os resultados obtidos foram organizados e subdivididos nos seguintes temas: aprendizagem, motivação e colaboração.

##### 5.1 Aprendizagem

A aprendizagem foi testada com apenas uma turma de 20 crianças ( $N=20$ ). Para tal, foi realizado um pré-teste no primeiro dia, no segundo dia as crianças brincaram com o jogo e só no terceiro dia, realizaram o pós-teste. Os testes consistiam numa tabela onde num lado tinha os animais e no cabeçalho os diferentes meios. O objectivo consistia em colocar um X no correspondente meio onde habitam os respectivos animais.

A principal métrica utilizada foi a diferença dos resultados entre os pré e os pós-testes, tendo em consideração o número de respostas erradas num conjunto de 8 perguntas. A diferença do número médio de respostas erradas no pré-teste foi maior do que no pós-teste.

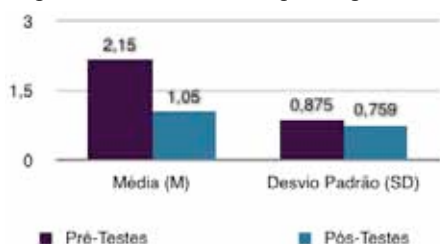


Gráfico 1: Resultados do número médio de respostas erradas nos pré e pós-testes.

Esta diferença mostrou ser estatisticamente extremamente significativa com ( $t(38)=4.25$ ,  $p=0.0001$ ).

Verificamos que o número de respostas erradas diminuiu, o que sugere que o jogo poderá ajudar as crianças na aprendizagem de determinados conceitos. Contudo, no pré-escolar, as competências a adquirir pelas crianças são outras (comportamento em sociedade, desenvolvimento das capacidades motoras, colaboração), e por isso focou-se o âmbito desta investigação noutros aspectos.

##### 5.2 Motivação

A motivação foi essencialmente medida através de observações, gravações de vídeo e análise dos questionários feitos às 101 crianças. Os questionários foram elaborados com ajuda das educadoras.

A motivação mostrou-se evidente quando as crianças escolheram ficar a brincar com o jogo em vez de ir



Figura 3: Crianças na escola a interagir com o sistema.

brincar para o recreio e por nunca terem desistido do jogo até conseguir chegar à solução. Outra forma de analisar a motivação foi através das próprias expressões, quer faciais (ex: sorrisos, espanto, alegria), verbais (ex: “UAUU”, “Que fixe!”) e corporais (ex: dar saltinhos e palmas) que as crianças faziam enquanto jogavam. De facto existe um incentivo de descoberta, uma vez que o jogo apela à curiosidade das crianças porque os marcadores são vistos apenas um de cada vez. A Figura 4 mostra algumas das muitas fotografias que evidenciam os grandes níveis de motivação alcançados.



Figura 4: A motivação das crianças enquanto interagem com o sistema.



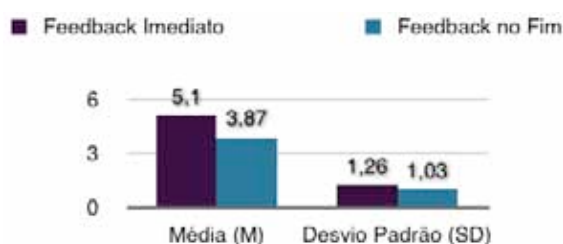
### 5.3 Colaboração

Os níveis de colaboração foram medidos sobre diferentes condições: utilizando feedback imediato versus feedback apenas quando todo o tabuleiro está completo.

A questão colocada foi: Será que a possibilidade de fornecer feedback imediato aumenta os níveis de colaboração entre as crianças enquanto brincam ao jogo?

Para testar esta hipótese, foram gravados os tempos de conclusão dos jogos, e analisamos as gravações de todos os testes com as 101 crianças, de forma a medir a variável dependente (colaboração) da seguinte forma quantitativa: (i) número de comentários colaborativos feitos pelas crianças, (ii) número de correcções feitas pelas crianças, incluindo gestos como apontar, e o (iii) número de tentativas feitas até chegar a uma possível solução.

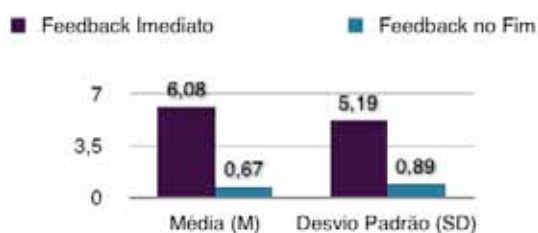
Os tempos de conclusão (em minutos) para a versão do feedback imediato ( $N=13$ , correspondendo a 13 grupos de crianças) foram maiores do que para a versão onde o feedback é dado apenas quando o tabuleiro está completo ( $N=12$ ).



**Gráfico 2: Resultados dos tempos de conclusão para as diferentes versões do jogo.**

Esta diferença foi considerada estatisticamente significativa com ( $t(23)=2.66$ ,  $p=0.014$ ). Embora o tempo de conclusão não seja uma medida de colaboração, os resultados obtidos das interacções colaborativas sugerem que o maior tempo de conclusão para a versão do feedback imediato foi causado por um maior número de interacções colaborativas.

Quanto ao número de comentários colaborativos feitos, observou-se que estes também foram mais altos na versão do feedback imediato ( $N=13$ ) do que na outra versão ( $N=12$ ).



**Gráfico 3: Resultados do número de comentários colaborativos nas diferentes versões do jogo.**

Esta diferença foi considerada estatisticamente muito significativa com ( $t(23)=3.58$ ,  $p=0.0016$ ). As crianças

fizeram significativamente mais comentários ao jogar com a versão do feedback imediato.

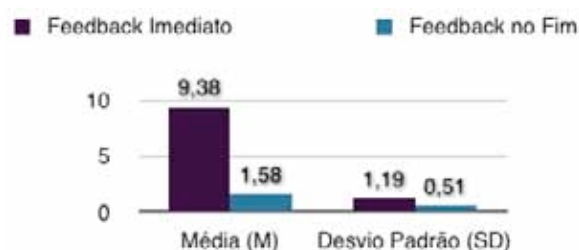
Quanto ao número de acções correctivas (por exemplo, mudar um marcador já colocado no tabuleiro para outra posição e gestos tal como apontar), observou-se novamente um número maior na versão do feedback imediato ( $N=13$ ) do que na outra versão ( $N=12$ ).



**Gráfico 4: Resultados do número de acções correctivas nas diferentes versões do jogo.**

Podemos verificar novamente, que esta diferença foi estatisticamente muito significativa com ( $t(23)=3.60$ ,  $p=0.0015$ ).

Finalmente, o número de tentativas feitas até chegar à solução foi muito maior na versão do feedback imediato ( $N=13$ ) do que na outra versão ( $N=12$ ).



**Gráfico 5: Resultados do número de tentativas feitas até chegar a uma solução, nas diferentes versões do jogo.**

Neste caso esta diferença mostrou ser estatisticamente extremamente significativa, com ( $t(23)=20.97$ ,  $p<0.0001$ ).

Estes valores sugerem que o feedback imediato poderá desempenhar um papel importante no aumento do número de comportamentos colaborativos e interacções entre as crianças do pré-escolar. Além disso, e de acordo com o feedback recebido das educadoras no final dos testes, as crianças efectivamente colaboraram umas com as outras, ajudando-se mutuamente através de comentários, correcções e brincando de uma forma ordenada.

Dentro da colaboração também foram feitos estudos em relação ao sexo e idades. No estudo em questão estiveram envolvidas as 101 crianças que participaram nos testes.

Verificamos que, no geral, o número de vezes que as crianças fizeram correcções físicas, tais como mudar uma peça de lugar ou simplesmente apontar, foi superior ao número de vezes que fizeram comentários.

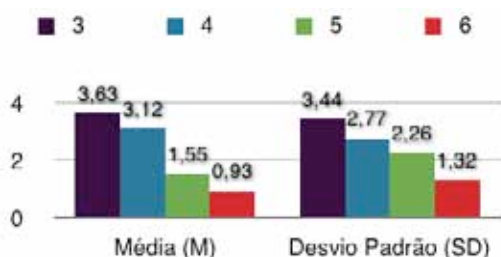


**Gráfico 6: Resultados do número de correcções físicas e verbais.**

Esta diferença mostrou ser estatisticamente muito significativa com ( $t(200)=2.93, p=0.0037$ ). O que mostra que as crianças tendem a fazer mais acções físicas do que comentários verbais.

O estudo por idades teve como base: 24 crianças de 3 anos, 33 crianças de 4 anos, 31 crianças de 5 anos e finalmente 13 crianças de 6 anos.

Para esta amostra obtivemos os seguintes resultados correspondentes ao total das acções colaborativas, quer físicas, quer verbais:



**Gráfico 7: Resultados do total de acções colaborativas por idades.**

As diferenças entre os resultados das crianças de 3 e 4 anos com ( $t(55)=0.62, p=0.538$ ), bem como das crianças de 5 e 6 anos com ( $t(42)=0.922, p=0.3631$ ) não mostraram ser estatisticamente significativas.

O mesmo não se verifica com os resultados obtidos entre as crianças de 3 e 6 anos com ( $t(35)=2.71, p=0.0104$ ), bem como entre as crianças de 4 e 5 anos com ( $t(62)=2.48, p=0.016$ ), cujas diferenças mostraram ser estatisticamente significativas.

Por fim, os resultados entre as crianças de 3 e 5 anos com ( $t(53)=2.70, p=0.0093$ ) e entre as crianças de 4 e 6 anos com ( $t(44)=2.72, p=0.0094$ ), foram os que mais se evidenciaram apresentando diferenças estatisticamente muito significativas. Concluimos que as crianças mais novas foram as que mais colaboraram no jogo.

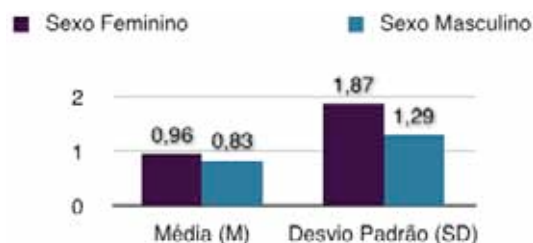
O estudo feito por sexos teve como base 50 crianças do sexo feminino e 51 crianças do sexo masculino. Comparando os resultados obtidos do número de acções físicas colaborativas femininas com o número de acções físicas colaborativas masculinas, verificamos que os rapazes foram aqueles que fizeram mais acções físicas.



**Gráfico 8: Resultados do número de acções físicas por sexos.**

No entanto, esta diferença foi considerada estatisticamente pouco significativa com ( $t(99)=1.87, p=0.0638$ ).

Os resultados obtidos para o número de comentários femininos foram levemente superiores aos masculinos, mas não foram estatisticamente significativos com ( $t(99)=0.41, p=0.6846$ ).



**Gráfico 9: Resultados do número de comentários por sexos.**

No entanto, verificamos que as raparigas tendem a usar mais as expressões verbais enquanto que os rapazes tendem a usar mais as acções físicas. O que demonstra que é bom fazer grupos mistos.

Os resultados obtidos do número total de acções colaborativas femininas em comparação com as masculinas foram muito parecidos, não sendo estatisticamente significativos com ( $t(99)=0.0001, p=0.9999$ ).



**Gráfico 10: Resultados do número total de acções colaborativas por sexos.**

Podemos observar, no entanto, que os rapazes foram aqueles que apresentaram um maior número de acções colaborativas.

## 6. CONCLUSÕES

Após este estudo, podemos concluir que existem princípios de design extremamente relevantes para sistemas direccionados a crianças com idades compreendidas entre

os 3 e os 5/6 anos, e a maioria desses princípios são obtidos por meio das educadoras, das práticas de observação e sessões de avaliação como as que foram feitas neste estudo.

Verificamos que todas as crianças, incluindo as mais novas, mostraram facilidades de coordenação ao combinar as duas condições necessárias para o funcionamento do jogo, ou seja, colocar o marcador sob a câmara e olhar para o ecrã ou para a projecção, onde apareciam os objectos em 3D e as respostas.

## 6.1 Aprendizagem

Em termos de aprendizagem, concluímos que através do brincar com o jogo, o número de respostas erradas diminuiu significativamente, o que sugere que o jogo poderá ajudar as crianças a aprender simples conceitos.

Uma vez que estas crianças perdem frequentemente o foco de atenção, especialmente com um jogo, tínhamos que este pudesse prejudicar o processo de aprendizagem, mas os resultados obtidos sugerem que o jogo de forma alguma prejudica o processo de aprendizagem.

De acordo com o feedback das educadoras, o jogo parece ser uma forma promissora para complementar os métodos tradicionais de ensino.

## 6.2 Motivação

Observamos níveis altos de motivação enquanto as crianças brincavam com o jogo. Esta afirmação deve-se ao facto de nenhuma criança ter desistido do jogo antes de alcançar a solução e também devido às expressões faciais, verbais e corporais que faziam ao longo do jogo mostrando uma grande satisfação. Os resultados dos questionários mostraram que a maior parte das crianças gostou muito do jogo (96%) e todas, sem excepção, gostariam de jogar novamente ao jogo. O facto das crianças ficarem muito envolvidas na actividade e por conseguirem atingir grandes níveis de concentração, verificados pelas observações feitas, mostrou que estavam motivadas com o jogo. O facto de optarem por ficar a jogar ao jogo dentro da sala (no caso do teste na escola) em vez de ir brincar para o recreio, também mostrou a motivação das crianças pelo jogo.

O uso do projector, que contribuiu com uma imagem grande do jogo projectada na parede ou na tela, foi também um contributo para a motivação das crianças, uma vez que estas gostam de coisas diferentes e que chamem a atenção. O projector foi um factor crucial de design, pois permitiu que toda a turma participasse e colaborasse no jogo. No geral, grandes projecções de imagem e uma boa combinação das peças físicas do jogo com elementos virtuais, podem ser eficazes para aumentar a motivação e os níveis de colaboração entre as crianças.

A utilização das cores e das frases sonoras para o feedback, foram muito bem aceites pelas crianças. Estas compreenderam logo os seus significados e mostraram satisfação.

Manter as crianças motivadas durante todo o jogo é uma questão fundamental de design.

## 6.3 Colaboração

O foco deste estudo baseou-se em promover colaboração. Analisamos diversas variáveis tais como o número de comentários feitos pelas crianças, o número de correcções, incluindo os gestos de apontar e o número de tentativas até alcançar uma possível solução.



Figura 5: Colaboração entre as crianças enquanto jogam.

Os resultados sugerem que o feedback imediato teve um papel muito importante, aumentando o número de comportamentos colaborativos e interacções entre as crianças.

Com este estudo verificamos um equilíbrio no total de acções colaborativas realizadas, quer físicas quer verbais, entre os diferentes sexos. No entanto as crianças tendem a fazer mais acções físicas (tais como corrigir um marcador já colocado no tabuleiro ou simplesmente apontar) do que acções verbais. Verificamos ainda que as crianças de 3 e 4 anos foram as que mais colaboraram. No geral, as crianças mais novas colaboraram mais do que as mais velhas. Estes resultados poderão estar relacionados com níveis superiores de curiosidade por parte das crianças mais novas, pelo facto destas possuírem menos conhecimentos que as mais velhas ou ainda devido ao facto destas poderem ter uma maior liberdade de expressão comparando com as mais velhas, que poderá advir da forma de trabalhar das próprias educadoras.

Em termos da forma de como a colaboração é feita, verificamos que novamente os resultados são muito parecidos. No entanto, os rapazes foram os que apresentaram um maior número de colaborações físicas e as raparigas apresentaram um maior número de colaborações verbais. Assim, compreende-se que a formação de grupos com ambos os sexos é um aspecto positivo a ter em conta na realização dos jogos.

De acordo com o feedback das professoras, as crianças de facto colaboraram, ajudando-se umas às outras. Assim, o jogo de tabuleiro de Realidade Aumentada actua como uma plataforma de comunicação entre as crianças.

Face ao exposto, compreende-se o potencial do jogo em Realidade Aumentada no desenvolvimento da aprendizagem, motivação e colaboração entre os alunos.

## 7. AGRADECIMENTOS

Os autores gostariam de agradecer às educadoras Clara Cristina Pessanha, Marta Costa, Ilda Santos, Helena Santos, Maria Franco, Rafaela Calaça, Marília Santos, Maria Caeiro, e Sílvia Carreira, pelo trabalho desenvolvido durante o design colaborativo do jogo.

## 8. REFERÊNCIAS

- [Africano04] Africano, D., Berg, S., Lindbergh, K., Lundholm, P., Nilbrink, F., and Persson, A. (2004). Designing tangible interfaces for children's collaboration. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, CHI '04. ACM, New York, NY, 853-868.
- [Amante04] Amante, L. Explorando as novas tecnologias em contexto de educação pré-escolar: A actividade de escrita. *Aná. Psicológica*. [on-line]. mar. 2004, vol.22, no.1 [citado 31 Outubro 2009], p.139-154, ISSN 0870-8231.
- [Belotti09] Bellotti, F., Berta, R., Gloria, A. D., and Primavera, L. (2009). Enhancing the educational value of video games. *Computers in Entertainment*, 7, 2 (Jun. 2009), pp. 1-18.
- [Duarte05] Duarte, M., Cardoso, A. & Lamounier Jr., E. (2005). Using Augmented Reality for Teaching Physics. *WRA '2005 - II Workshop on Augmented Reality*, pp. 1-4.
- [Gibson03] Gibson, J. P. (2003). A noughts and crosses Java applet to teach programming to primary school children. In *Proceedings of the 2nd international Conference on Principles and Practice of Programming in Java*, PPPJ, vol. 42. Computer Science Press, New York, NY, pp.85-88.
- [Hsieh08] Hsieh, Min-Chai. e Lee, Jiann-Shu. (2008). AR Marker Capacity Increasing for Kindergarten English Learning. National University of Tainan, Hong Kong.
- [Inkpen97] Inkpen, K. M., Booth, K. S., Klawe, M., & McGrenere, J. (1997). The Effect of Turn-Taking Protocols on Children's Learning in Mouse-Driven Collaborative Environments. In *Proceedings of Graphics Interface* (GI 97). Canadian Information Processing Society, pp. 138-145.
- [Kannetis09] Kannetis, T. and Potamianos, A. (2009). Towards adapting fantasy, curiosity and challenge in multimodal dialogue systems for preschoolers. In *Proceedings of the 2009 International Conference on Multimodal interfaces*, ICMI-MLMI'09. ACM, New York, NY, 39-46.
- [Kaufmann02] Kaufmann, H. and Schmalstieg, D. (2002). Mathematics and geometry education with collaborative augmented reality. In *ACM SIG-GRAPH 2002 Conference Abstracts and Applications*, ACM, New York, NY, pp. 37-41.
- [Mayo07] Mayo, M. J. (2007). Games for science and engineering education. *Communications of the ACM*, 50, 7 (Jul.2007), pp. 30-35.
- [Medicherla10] Medicherla, P. S., Chang, G., and Moreale, P. (2010). Visualization for increased understanding and learning using augmented reality. In *Proceedings of the international Conference on Multimedia information Retrieval*, MIR'10. ACM, New York, NY, pp. 441-444.
- [Papert96] Papert, S. (1996). *The Connected Family: Bridging the Digital Generation Gap*. Longstreet Press.
- [Salomon90] Salomon, G. (1990). Studying the flute and orchestra: Controlled experimentation vs. Whole classroom research on computers. *International Journal of Educational Research*, 14, pp. 37-47.
- [Schrier06] Schrier, K. (2006). Using augmented reality games to teach 21st century skills. In *International Conference on Computer Graphics and Interactive Techniques*, ACM SIGGRAPH 2006 Educators program.
- [Shelton02] Shelton, B. and Hedley, N. (2002). Using Augmented Reality for Teaching Earth-Sun Relationships to Undergraduate Geography Students. In *The First IEEE International Augmented Reality Toolkit Workshop*, Darmstadt, Germany, September 2002, IEEE Catalog Number: 02EX632 ISBN: 0-7803-7680-3.
- [Stewart98] Stewart, J., Raybourn, E. M., Bederson, B., & Druin, A. (1998). When two hands are better than one: Enhancing collaboration using single display groupware. In *Proceedings of Extended Abstracts of Human Factors in Computing Systems* (CHI '98).
- [Sutton86] Sutton-Smith, B. *Toys as culture*. New York: Gardner Press, 1986.
- [Tettegah06] Tettegah, S. & Taylor, K. & Whang, E. & Meistninkas, S. & Chamot, R. (2006). Can virtual reality simulations be used as a research tool to study empathy, problems solving and perspective taking of educators?: theory, method and application. *International Conference on Computer Graphics and Interactive Techniques*, ACM SIGGRAPH 2006 Educators program, Article No. 35.



# Rumo a Interfaces Tácteis Acessíveis

Tiago Guerreiro, Hugo Nicolau, Joaquim Jorge, Daniel Gonçalves

IST / UTL / INESC-ID

Rua Alves Redol, 9, 1000-029 Lisboa

{tjvg, hman}@vimmi.inesc-id.pt, {jaj, daniel.goncalves}@inesc-id.pt

---

## Sumário

*Este artigo apresenta uma avaliação efectuada a 15 utilizadores tetraplégicos com o objectivo de compreender as suas capacidades num conjunto de técnicas de interacção (Tapping, Crossing, Exiting e Gestos Direccionais) e respectivas parametrizações (posição, tamanho e direcção). Os resultados mostraram que para cada técnica a eficácia e precisão varia de acordo com as diferentes parametrizações. Regra geral, o Tapping (método tradicional) foi a técnica de interacção preferida e entre as mais eficazes. Isto mostra que é possível criar interfaces unificadas e acessíveis a utilizadores com e sem deficiências, caso existam métodos de parametrização ou adaptação apropriados.*

## Palavras-chave

*Ecrã Táctil, Tetraplégico, Técnicas de Interacção, Avaliação, Dispositivos Móveis.*

---

## 1. INTRODUÇÃO

Ao longo dos últimos anos os tradicionais teclados dos dispositivos móveis têm vindo a ser substituídos por ecrãs tácteis. Este tipo de interfaces oferece diversas vantagens em relação aos seus homólogos. Em particular, são mais flexíveis, permitindo mostrar diferentes interfaces no mesmo dispositivo (e.g. teclado de 12 teclas, QWERTY) ou adaptar-se às preferências de cada utilizador. Mais, a capacidade de tocar e manipular directamente os seus elementos, sem dispositivos intermediários, revela ser uma forma natural e simples de interacção.

No entanto, as interfaces tácteis apresentam também alguns desafios: não oferecem pistas tácteis nem a estabilidade física dos teclados, tornando-se difícil seleccionar alvos no ecrã. Estas dificuldades tornam-se especialmente relevantes para indivíduos com baixa precisão ou controlo motor, como é o caso de utilizadores tetraplégicos.

Por outro lado, os deficientes motores podem beneficiar imenso do uso de ecrãs tácteis, sendo que muitos não têm a força ou destreza para pressionar botões físicos [Myers02]. Mais, o uso de PDAs já mostrou ser uma alternativa viável aos periféricos tradicionais (i.e. rato e teclado), permitindo que a mesma interface fosse usada em diferentes locais e contextos.

O seu elevado grau de flexibilidade torna estes dispositivos ideais para soluções personalizadas ou adaptativas, capazes de ir ao encontro das necessidades de cada utilizador. Isto apresenta grandes oportunidades para esta população, particularmente àqueles que têm falta de força ou controlo nos seus membros superiores. No entanto, é ainda muito difícil desenhar interfaces tácteis para deficientes motores, uma vez que não existe um conhecimento empírico das vantagens e limitações de cada técnica de

interacção, considerando o perfil dos utilizadores. As suas necessidades e capacidades têm de ser tidas em conta de forma a podermos desenvolver interfaces, para dispositivos móveis, mais precisas e eficazes.

Para superar esta lacuna efectuámos avaliações com 15 utilizadores tetraplégicos (Figura 1), testando 4 técnicas de interacção distintas. Estas incluíam o *Tapping*, *Crossing*, *Exiting* e *Gestos Direccionais*. Explorámos também as características do ecrã e parâmetros da interface em busca de guias de desenho que possam tornar a experiência de utilização mais agradável.

O nosso objectivo principal foi fornecer o conhecimento empírico para um desenho de interfaces tácteis acessíveis. Focamos a nossa atenção no tamanho dos alvos e posições do ecrã que são associadas a um aumento de desempenho (i.e. cantos e arestas). Neste artigo descrevemos toda a fase de avaliação, apresentamos e analisamos os resultados obtidos, não só para cada técnica, mas também entre técnicas. Por fim, retiramos algumas conclusões e propostas de desenho.



**Figura 1 – Participante a interagir com o ecrã táctil durante uma sessão de avaliação.**

## 2. TRABALHO RELACIONADO

Alguns trabalhos tentaram já melhorar o acesso a dispositivos móveis com ecrãs tácteis. Wobbrock et al. [Wobbrock03] propuseram o uso de uma caneta juntamente com a utilização dos cantos e arestas do dispositivo para a escrita de texto. Os resultados mostraram que o *EdgeWrite* oferece aos utilizadores, com deficiências motoras, um elevado grau de estabilidade nos seus movimentos. Usando uma abordagem semelhante, o *Barrier Pointing* [Froehlich07] usa as arestas e cantos do ecrã para melhorar a precisão de tarefas de selecção. Executando gestos em direcção às barreiras do dispositivo os utilizadores conseguiram seleccionar alvos com uma maior estabilidade física.

Apesar de estes trabalhos explorarem as propriedades físicas dos dispositivos para auxiliar a interacção entre deficientes motores e ecrãs tácteis, existe ainda pouco conhecimento empírico acerca do seu desempenho com outras técnicas. No entanto, muitos trabalhos de investigação têm focado a sua atenção em tentar perceber e maximizar o desempenho dos utilizadores, sem deficiências, neste tipo de dispositivos [Lee09, Lin07, Mizobuchi05, Parhi06, Park08, Perry08].

O tamanho dos alvos é um dos principais problemas no estudo de interfaces tácteis. Parhi et al. [Parhi06] conduziu um estudo para determinar o tamanho óptimo de alvos para a utilização do polegar em dispositivos móveis. Os resultados mostraram que tamanhos entre 9.2 mm e 9.6 mm podem ser usados sem uma perda de desempenho. De forma semelhante, Park et al. [Park08] analisou três tamanhos de teclas diferentes. Os resultados revelaram que o tamanho maior (10 mm) obteve o melhor desempenho e satisfação dos utilizadores. Lee e Zhai [Lee09] obtiveram resultados semelhantes, em que alvos com tamanho inferior a 10 mm sofreram uma diminuição de desempenho.

Considerando a posição dos alvos, os utilizadores preferem-nos no centro do ecrã, visto serem mais fáceis e confortáveis de alcançar. No entanto, as melhores taxas de eficácia ocorrem nos alvos perto das arestas [Park08].

Os trabalhos anteriores sugerem algumas recomendações em relação ao tamanho e localização dos alvos para interfaces móveis. Mizobuchi et al. [Mizobuchi05] efectuou um estudo para determinar como a escrita de texto num ecrã táctil, usando uma caneta, é afectada pelo movimento do utilizador. Os autores sugerem que um teclado virtual não pode ter teclas com tamanho inferior a 3 mm, para obter uma taxa de erro de 2%. No entanto, situações de deslocamento mais exigentes podem requerer alvos maiores [Lin07]. Os resultados mostram que utilizadores a andar num percurso de obstáculos conseguem seleccionar alvos de 6.4 mm com uma taxa de erro de 10%.

Apesar de estes estudos terem sido efectuados a utilizadores com deficiências situacionais [Sears03] e usarem caneta, podem ainda assim fornecer conhecimentos válidos no desenho de interfaces para deficientes motores. De facto, estes utilizadores podem experienciar problemas semelhantes, como o tremor e instabilidade física.

No entanto, estas aparentes semelhanças não são suficientes para generalizar os resultados.

Como podemos verificar, existe uma carência de conhecimento referente ao uso de superfícies tácteis por parte de deficientes motores. A experiência descrita neste artigo tenta colmatar esta lacuna testando várias técnicas de interacção, características e parametrizações, fornecendo um suporte empírico para o desenho informado de interfaces tácteis.

## 3. AVALIANDO AS TÉCNICAS DE INTERACÇÃO

Os dispositivos com ecrãs tácteis colocam ambos desafios e oportunidades à comunidade científica. De facto, apesar de estas interfaces imporem uma menor exigência motora, são menos exploradas que as interfaces baseadas em botões. Mais, para nosso conhecimento, estudos de desempenho de deficientes motores com ecrãs tácteis, ainda não foram realizados. Em particular, com utilizadores tetraplégicos que possuam algum movimento residual nos seus membros superiores.

### 3.1 Técnicas de Interacção e Variações

Nesta experiência escolhemos um conjunto de técnicas representativas das diferentes formas de interacção aplicadas a superfícies tácteis [Froehlich07, Wobbrock03]. Em seguida realizamos um estudo com utilizadores tetraplégicos observando o seu desempenho em dispositivos móveis com este tipo de interfaces.

Considerámos dois métodos de interacção base: *tapping* e gestos. Ao realizar um gesto, os utilizadores poderiam: cruzar o alvo; confirmar a selecção saindo do ecrã através do alvo; ou usar gestos direccionais (Figura 2).

**Tapping** no ecrã consiste em seleccionar o alvo através do toque (i.e. pousar o dedo no alvo). Esta é a técnica de interacção mais comum nas aplicações actuais, devido à sua facilidade de utilização. Nesta técnica, os alvos foram apresentados em 3 tamanhos diferentes (7, 12 e 17 mm) em todas as posições do ecrã: cantos, arestas e centro. Os tamanhos foram escolhidos com base em estudos anteriores com utilizadores sem deficiências [Lee09, Parhi06, Park08, Perry08].

**Crossing**, ao contrário do *Tapping* não envolve posicionar o dedo dentro do alvo. Ao invés, este é seleccionado cruzando-o. Trabalhos anteriores demonstraram que esta técnica oferece um melhor desempenho, a deficientes motores, do que as técnicas de interacção tradicionais [Wobbrock08]. Os alvos foram apresentados apenas nas posições centrais do ecrã (ver Figura 3) em 3 tamanhos distintos.

**Exiting** é semelhante ao *Crossing*, mas a selecção consiste na execução de um gesto em direcção à borda do ecrã,

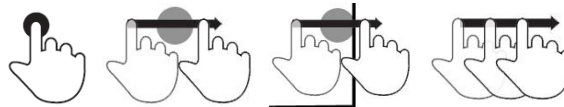


Figura 2 - Tapping, Crossing, Exiting, e Gestos Direccionais.





**Figura 3 - Qtek 9000. Zonas do ecrã (esquerda): preto – cantos; cinza – arestas; branco – centro; Alvo (direita).**

saindo pelo alvo pretendido. Para esta técnica, os alvos foram mostrados em todos os tamanhos, no entanto, apenas nos cantos e arestas.

Os *Gestos Direccionais* eram a única técnica de interacção que não consistia na selecção de alvos. Os utilizadores podiam executar gestos direccionais em qualquer zona do ecrã. Esta técnica foi escolhida devido à sua flexibilidade e, assim como o *Tapping*, é bastante usada nas interfaces tácteis actuais.

### 3.2 Questões de Investigação

Esta avaliação tenciona responder a algumas questões relativamente à nossa população-alvo:

1. Qual o melhor tamanho para cada técnica?
2. Qual a melhor zona do ecrã (canto, aresta ou centro) para cada técnica de interacção?
3. Qual a melhor técnica de interacção?
4. Qual a melhor técnica para cada zona do ecrã?
5. Qual a melhor combinação técnica-tamanho-zona?

### 3.3 Participantes

Foram recrutados quinze participantes num centro de fisioterapia (13 do sexo masculino), com idades compreendidas entre 28 e 64 anos e lesões de C4 a C7. Todos os utilizadores possuíam algum controlo residual nos membros superiores (9 com a mão direita como dominante).

Antes da experiência todos os participantes executaram uma avaliação de capacidades (aperto, alcance e precisão). Esta avaliação funcional teve como objectivo ter uma medida objectiva das capacidades dos participantes, complementando o nível de lesão. No entanto, não foram encontradas correlações entre as suas características e desempenho nas tarefas propostas [Guerreiro10].

Em relação à familiaridade com tecnologia, todos os participantes possuíam um telefone móvel e usavam-no diariamente. No entanto, nenhum deles tinha um dispositivo com ecrã táctil.

### 3.4 Material

Nesta experiência usámos um PDA, QTEK 9000 (Figura 3), com *Windows Mobile 5.0*. O dispositivo móvel tinha

uma resolução de 640x480 (73x55 mm) pixels e possuía bordas salientes. Durante a avaliação foi feita captura de vídeo e todas as interacções com o dispositivo foram guardadas através de um mecanismo de *logging* para posterior análise.

### 3.5 Procedimento

No início da avaliação foi dito aos participantes que o objectivo geral do estudo era investigar e comparar diferentes técnicas de interacção para ecrãs tácteis e a sua adequação às capacidades de utilizadores tetraplégicos. Em seguida, foi feito um questionário e o teste de avaliação das capacidades de forma a caracterizar cada participante. Por fim, foi-lhes explicada e demonstrada cada técnica de interacção (*Tapping*, *Crossing*, *Exiting* e *Gestos Direccionais*).

Para reduzir efeitos de aprendizagem os participantes tiveram um período de treino antes de avaliar cada uma das técnicas, onde podiam mover o dispositivo e colocá-lo na posição mais confortável. Todas as sessões foram realizadas num local calmo (em suas casas ou no centro de fisioterapia), sempre com os participantes sentados nas suas cadeiras e com uma mesa de apoio, ou descanso para o braço, à sua frente (Figura 1). Os participantes eram livres de usar qualquer parte da mão para interagir com o ecrã táctil.

A cada utilizador foi-lhe pedido para efectuar selecções de alvos com cada uma das técnicas (*Tapping*, *Crossing* e *Exiting*). Para os *Gestos Direccionais* não existiam alvos e os participantes tinham apenas de realizar gestos em diferentes direcções (e.g. norte). Existiam 16 possíveis direcções, incluindo diagonais e gestos com suporte das arestas (e.g. norte usando a aresta do lado direito).

Os participantes tinham apenas uma tentativa para completar a tarefa e não eram informados do seu resultado. No entanto, recebiam retorno de que uma acção tinha sido realizada. Após 2 segundos o alvo seguinte aparecia no ecrã.

Para as técnicas de selecção de alvos foram usados 3 tamanhos diferentes (7, 12 e 17 mm). Estes tamanhos foram escolhidos com base na literatura existente [Lee09, Parhi06, Park08, Perry08]. Os testes foram seleccionados de forma aleatória. Em cada conjunto técnica-tamanho, as posições dos alvos foram também apresentadas aleatoriamente. Assim, cada participante realizou 142 acções, ou seja, um total 2130 acções em toda a experiência.

### 3.6 Medidas

As medidas usadas nesta avaliação foram obtidas através da nossa aplicação de teste, responsável por capturar todas as interacções com o dispositivo. As variáveis dependentes eram *Erros de Tarefa*, *Precisão*, *Erro de Movimento* e *Tempo de Movimento*.

Para as técnicas de selecção de alvos (*Tapping*, *Crossing* e *Exiting*), a *Precisão* foi calculada com base na distância mínima ao centro do alvo. Para os *Gestos Direccionais*, a *Precisão* correspondia à distância média ao eixo da direcção pretendida.

Para as abordagens gestuais (*Crossing*, *Exiting* e *Gestos Direccionais*) os tempos e erros [Soukoreff04] dos movimentos foram também capturados. O *Tempo de Movimento* correspondia à duração do toque no ecrã durante a realização do gesto. O *Erro de Movimento* era a medida absoluta do desvio ao eixo do gesto realizado. A diferença entre *Erro de Movimento* e *Precisão* é que o primeiro consiste numa medida de estabilidade do movimento realizado, enquanto o último tem em conta apenas o objectivo da tarefa (direcção ou proximidade ao alvo).

### 3.7 Desenho da Experiência e Análise

Nesta avaliação variámos a técnica de interacção, tamanho e posição dos alvos. Mais, usamos um desenho intra-grupos, onde cada participante testou todas as condições. Para a análise da posição criamos dois factores adicionais: Áreas Verticais e Distância (ver Figura 6). A última reflecte a posição do alvo em relação ao suporte do participante (o nível 1 refere-se à posição mais próxima e o nível 5 à mais distante). Apesar de estar visualmente representada para participantes dextros, esta medida foi normalizada para ambas as mãos.

Testes de *Shapiro-Wilkinson* sobre os valores observados para *Erros de Tarefa*, *Precisão*, *Erro* e *Tempo de Movimento* mostraram seguir uma distribuição normal para todas as técnicas de interacção. Logo, foi aplicada uma *Repeated-measures ANOVA* na análise dos resultados.

## 4. RESULTADOS

O nosso principal objectivo consiste em perceber as capacidades e limitações da população-alvo no uso de ecrãs tácteis. Pretendemos assim estudar diferentes técnicas e realçar as suas vantagens e falhas. Primeiro, focamos em cada técnica separadamente; assim, se forem usadas numa interface, poderão sê-lo de forma sustentada. Depois, analisamos as técnicas em conjunto, comparando a sua eficácia em condições variadas. Este conhecimento permite a quem idealiza e desenha interfaces decidir como adequar a interacção em cada circunstância.

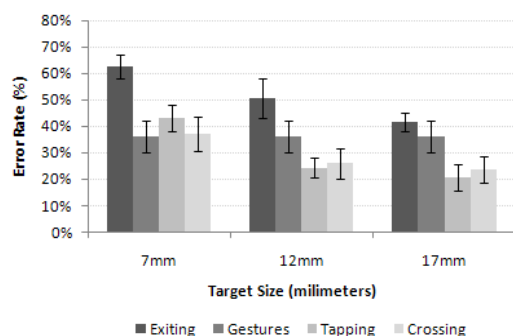


Figura 4 – Taxa de erros para cada Técnica e Tamanho. Barras de erro denotam IC de 95%.

### 4.1 Analisando cada técnica

As técnicas analisadas (resultados gerais de eficácia apresentados na Figura 4) têm diferentes essências e cada uma apresenta as suas vantagens e desvantagens. Nesta secção apresentamos os resultados obtidos para cada técnica e analisamos diferenças apenas entre variantes da mesma.

#### 4.1.1 Tapping

Esta técnica foi analisada em relação a *Erros* e *Precisão*. A Figura 5 apresenta as *Taxas de Erro* para cada *Posição* para todos os *Tamanhos*, dando uma visão geral sobre os resultados do *Tapping*. Olhando para variações da experiência:

**Tamanho.** Foi encontrado um efeito significativo do *Tamanho* nos *Erros* ( $F_{1,42}=25.10$ ,  $p<.001$ ). Um teste de comparações múltiplas mostrou diferenças significativas entre os tamanhos pequeno e médio, bem como entre os tamanhos pequeno e grande (Figura 4). Estes resultados sugerem 12 mm como um valor aproximado aceitável para aquisição de alvos por pessoas com deficiências motoras com resíduo nos membros superiores. Em relação à *Precisão*, não foram encontradas diferenças significativas entre tamanhos.

**Cantos.** Não encontramos diferenças significativas da *Posição* do alvo (canto ou centro) nos *Erros* (tamanhos médio e grande). Um efeito menor ( $F_{1,28}=3.10$ ,  $p<.1$ ) foi encontrado no tamanho menor (menos erros nos cantos). No que diz respeito à *Precisão*, foi encontrado um efeito significativo no tamanho médio, indicando melhores resultados nos cantos ( $F_{1,14}=8.941$ ,  $p<.01$ ). Isto indica que os cantos oferecem maior estabilidade para um movimento preciso embora não se reflecta numa maior eficácia.

**Arestas.** Não foram encontradas diferenças significativas nos *Erros*, independentemente do *Tamanho* e da *Posição* do alvo ser na aresta ou não. No entanto, foi encontrado um efeito significativo da *Posição* (aresta ou centro) na *Precisão* para o tamanho pequeno ( $F_{1,28}=14.41$ ,  $p<.01$ ), médio ( $F_{1,28}=6.85$ ,  $p<.005$ ) e grande ( $F_{1,28}=27.67$ ,  $p<.001$ ), mostrando maior precisão nas arestas. Tal como nos alvos dos cantos, a estabilidade física oferecida pelas arestas é também benéfica para uma selecção mais precisa.

**Áreas Verticais.** Foi encontrada uma diferença significativa nos tamanhos médio ( $F_{1,42}=3.59$ ,  $p<.05$ ) e grande ( $F_{1,42}=5.19$ ,  $p<.05$ ) (Figura 6). Um teste de comparações

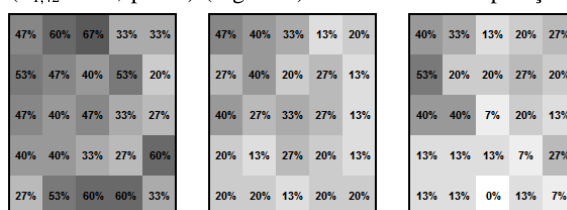
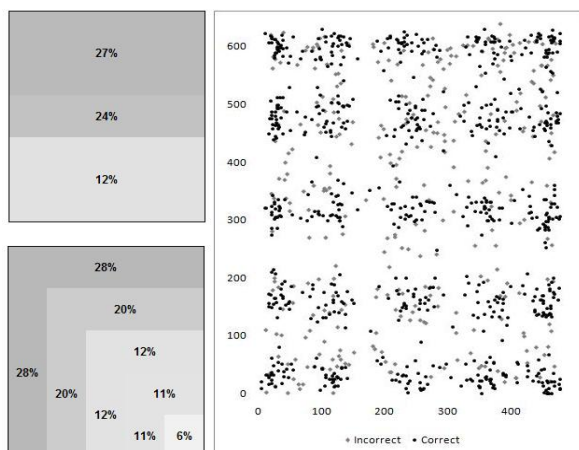


Figura 5 – Mapa de Taxas de Erros para cada Tamanho: 7mm (esquerda), 12 mm (centro) e 17mm (direita). Refere-se ao centro do alvo.



**Figura 6 – Taxa de Erros (esquerda): áreas verticais (cima) e níveis de distância (baixo) no tamanho maior. Taps realizados mostrando Erros e Precisão (direita).**

múltiplas mostrou diferenças entre os alvos de topo e os de baixo, com taxas de erro mais altas nos alvos do topo do ecrã. Relativamente à *Precisão*, um efeito menor foi encontrado nos tamanhos médio e grande, apontando também para diferenças entre alvos de topo e de baixo (mais precisão nos de baixo). Estes dados sugerem que os utilizadores seleccionam com maior eficácia e precisão alvos que estejam mais perto do ponto de fixação do seu braço.

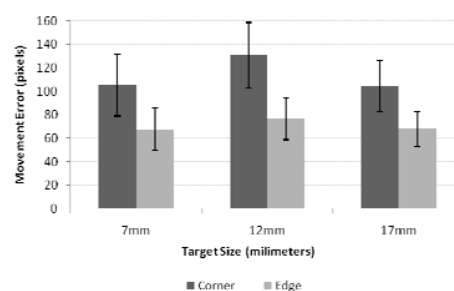
**Distância.** Similarmente ao resultado das áreas verticais, a *Posição (Distância)* (Figura 6) teve um efeito significativo nos *Erros* no tamanho grande ( $F_{1,70}=4.56, p<.01$ ). Testes *Post-hoc* mostraram diferenças significativas entre os níveis 1,2 (mais perto do braço do utilizador) e o nível 5. Existiu também um efeito significativo da *Posição (Distância)* na *Precisão*, no tamanho menor ( $F_{1,70}=6.04, p<.001$ ), entre os níveis 2,4 e o nível 5, e no tamanho médio ( $F_{1,70}=4.85, p<.01$ ) e tamanho grande ( $F_{1,70}=8.31, p<.001$ ) entre os níveis 2,3,4 e o nível 5.

#### 4.1.2 Crossing

O estudo desta técnica contemplou alvos nas nove posições centrais, evitando alvos junto às arestas/cantos (estes alvos foram avaliados separadamente como uma técnica particular – *Exiting*). A análise ao *Crossing* incluiu, para além dos *Erros* e *Precisão*, análise ao *Tempo* e *Erro de Movimento*.

**Tamanho.** Foi encontrado um efeito significativo do *Tamanho* dos alvo nos *Erros* cometidos ( $F_{1,42}=6.56, p<.01$ ). As diferenças significativas foram encontradas entre o tamanho mais pequeno e maior (Figura 4). Quanto à *Precisão*, um efeito menor foi detectado entre os mesmos tamanhos. Não foi encontrado qualquer efeito no *Tempo* ou *Erro de Movimento*. Esta ausência de efeitos sugere que o *Tamanho* não tem influência na forma como os utilizadores cruzam os alvos (o tipo de movimento e o tempo dispendido a executar o gesto).

**Áreas verticais.** Verticalmente, não foi encontrado nenhum efeito da *Posição (Áreas Verticais)* nos *Erros*,



**Figura 7 – Erro de Movimento, por tamanho, comparando cantos e arestas (Exiting). Barras de erro denotam um IC de 95%.**

*Precisão, Tempo ou Erro do Movimento.* Este resultado era esperado visto que todos os alvos se encontram em posições centrais, minimizando as diferenças verticais.

**Distância.** Não foram encontradas diferenças significativas da *Posição (Distância)* nos *Erros, Precisão, Tempo, ou Erro de Movimento*.

#### 4.1.3 Exiting

*Exiting* é semelhante ao *Crossing* mas é realizada com alvos junto às arestas e cantos. Assim, os utilizadores realizam o gesto de encontro à barreira física (limite do ecrã), saindo (*exiting*) do ecrã. A análise a esta técnica incluiu observação de *Erros, Precisão, Tempo e Erro de Movimento*.

**Tamanho.** Houve um efeito significativo do *Tamanho* nos *Erros* ( $F_{1,42}=7.77, p<.01$ ) entre o tamanho menor e maior. (Figura 4). Não foi encontrado qualquer efeito na *Precisão, Tempo ou Erro de Movimento*. Tal como no caso do *Crossing*, a ausência de significância sugere que o *Tamanho* não influencia a forma como o utilizador prepara e realiza o gesto.

**Cantos.** Não encontramos efeitos significativos da *Posição (cantos ou apenas aresta)* nos *Erros, Tempo ou Precisão*, em qualquer dos tamanhos. Quanto ao *Erro de Movimento*, existiram diferenças significativas nos tamanhos menor ( $F_{1,28}=5.04, p<.05$ ), médio ( $F_{1,28}=9.48, p<.01$ ) e maior ( $F_{1,28}=10.32, p<.01$ ) (Figura 7). Assim, os utilizadores têm semelhante eficácia nos cantos ou arestas mas produzem gestos mais erráticos quando o seu movimento é mais restringido (um alvo num canto tem menor superfície para o gesto do que um alvo na aresta). Observação dos movimentos mostrou que os utilizadores têm preferência por gestos específicos. Estas preferências podem colidir com as restrições impostas por um alvo no canto levando a um movimento convoluido.

**Áreas Verticais.** Houve um efeito significativo da *Posição* nos *Erros* no tamanho médio ( $F_{1,42}=3.42, p<.05$ ), mostrando maior eficácia nos alvos de topo. Também relevante, um efeito significativo foi encontrado na *Precisão* ( $F_{1,42}=4.43, p<.05$ ) e no *Erro de Movimento* ( $F_{1,42}=4.12, p<.05$ ), no tamanho pequeno, sugerindo que os utilizadores conseguiram contrapor a falta de eficácia com um movimento mais convoluido. No tamanho maior, não foram encontradas diferenças. Estes resultados contrastam com os resultados do *Tapping*. Análise

aos movimentos nas filmagens realizadas mostrou que para os alvos de topo, na maior parte dos casos bem sucedidos, o contacto com a superfície é realizado numa zona próxima do utilizador. Isto pode indicar que os utilizadores abordam melhor a superfície numa zona próxima mas conseguem depois produzir gestos eficazes em direcção ao topo (Figura 9 – esquerda e centro).

**Distância.** A *Posição (Distância)* mostrou ter um efeito significativo na *Precisão* no tamanho menor ( $F_{1,70}=4.08$ ,  $p<.05$ ). As diferenças mostraram-se significativas entre a distância intermédia (pior precisão) e ambos os alvos perto e longe. Observação dos testes sugere que os utilizadores seleccionam melhor alvos mais perto do suporte do braço (resultados apoiados pelos resultados do Tapping). Após o contacto, os gestos mais eficazes são aqueles com uma direcção mais rectilínea. Este facto também é suportado pelos resultados reportados para as Áreas Verticais. Um efeito significativo da *Posição (Distância)* foi também revelado no *Tempo*, tanto no tamanho médio ( $F_{1,70}=3.48$ ,  $p<.05$ ) como no maior ( $F_{1,70}=3.55$ ,  $p<.05$ ), sugerindo que quanto mais longe estão os alvos mais tempo demora a aquisição (implica gestos mais longos visto que o contacto é feito mais perto do suporte do braço).

#### 4.1.4 Gestos Direccionais

Nos gestos direccionais, não existem alvos nem consideramos tamanhos, apenas direcções. Este método inclui análise a *Erros*, *Precisão*, *Tempo* e *Erro de Movimento*.

**Arestas.** Não encontramos diferenças significativas entre *Posição* (gestos suportado pelas arestas ou em qualquer outra zona do ecrã) e *Erros* ou *Precisão*. No entanto, existiu um efeito significativo no *Erro de Movimento* ( $F_{1,28}=26.68$ ,  $p<.001$ ) (Figura 8). Este efeito foi encontrado entre gestos realizados nas arestas (menos erros) e realizados em qualquer zona central do ecrã. No entanto, estas diferenças são influenciadas pelos maus resultados obtidos nos gestos diagonais (que são significativamente piores que os gestos verticais e horizontais). Se estes gestos forem descartados, não existem diferenças significativas entre gestos realizados “no meio do ecrã” e suportados pelas arestas. Em relação ao *Tempo*, existe uma diferença significativa entre os gestos suportados pela aresta (mais rápidos) e os realizados no centro do ecrã ( $F_{1,70}=2.52$ ,  $p<.05$ ).

**Áreas Verticais.** Não encontramos qualquer efeito em nenhuma das variáveis dependentes.

**Direcção.** Não foram encontradas diferenças significativas entre a *Direcção* do gesto e os *Erros* ou *Precisão*. Muitos erros foram devido a toques não-intencionais no ecrã mas sem relação com qualquer direcção em particular. Inspecção visual sugeriu que algumas direcções são mais problemáticas, mas estas diferenças não foram significativas. Relativamente ao *Erro de Movimento* ( $F_{1,224}=15.14$ ,  $p<.001$ ) e *Tempo* ( $F_{1,224}=2.52$ ,  $p<.05$ ), as diferenças foram significativas. Os gestos diagonais são os mais erráticos, gestos suportados pelas arestas são

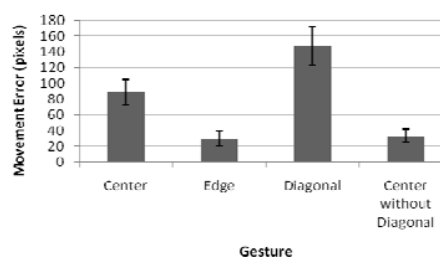


Figura 8 – *Erro de Movimento* considerando diferentes direcções e zonas de ecrã. Barras de erro denotam um IC de 95%.

mais lineares e gestos centrais (sem diagonais) situam-se entre os restantes.

## 4.2 Análise mista

A análise realizada a cada técnica reforça a ideia de que a eficácia e precisão são afectadas pelas características dos alvos, como a sua posição no ecrã ou tamanho. Este efeito tem proporções diferentes nas diferentes abordagens propostas. Já analisámos cada método individualmente. Iremos agora focar na comparação das técnicas e em perceber qual a mais adequada para certos pares posição/tamanho.

### 4.2.1 Uma visão geral sobre as técnicas

A apresenta as percentagens de *Erros* para todas as técnicas e tamanhos. Foi encontrado um efeito significativo da *Técnica* utilizada nos *Erros* no tamanho pequeno ( $F_{1,56}=5.97$ ,  $p<.01$ ), médio ( $F_{1,56}=5.66$ ,  $p<.01$ ) e grande ( $F_{1,56}=8.04$ ,  $p<.01$ ). Para o tamanho mais pequeno, as diferenças são significativas entre *Exiting* (pior) e todos os outros, enquanto para os outros tamanhos, o *Exiting* é significativamente pior que o *Tapping* e o *Crossing*. Em geral, o *Exiting* realça-se como um método menos eficaz embora as suas falhas diminuam quando comparado com abordagens por gestos direccionais, à medida que o tamanho do alvo aumenta. O *Tapping* e o *Crossing* revelaram-se como as técnicas mais eficazes sem diferenças significativas entre elas. Em suma, as diferenças nos *Erros* dos tamanhos médio para o grande não são significativas, sugerindo convergência e, assim, o tamanho médio como um bom compromisso.

### 4.2.2 Interagir com suporte da aresta

Pode-se argumentar que as bordas do ecrã oferecem um suporte positivo para a interacção. Nas técnicas consideradas, pedimos aos utilizadores para seleccionar alvos junto à aresta (*Tapping*) e para realizar gestos com suporte da mesma (*Gestos Direccionais*). Não foi encontrado um efeito da *Técnica* (*Tapping* na aresta vs. *Gestos* na aresta) nos *Erros* nos tamanhos pequeno e médio. Um feito menor foi encontrado no tamanho maior sugerindo maior eficácia em toques na aresta (*Tapping*) ( $F_{1,28}=3.15$ ,  $p<.1$ ). Isto é compreensível visto que as arestas forçam o utilizador a um determinado gesto numa determinada direcção (ex: da esquerda para a direita). Uma que pode ser impossível ou fácil de atingir. O *Tapping* é menos restritivo visto que o utilizador continua a poder abordar o alvo como mais lhe convier.

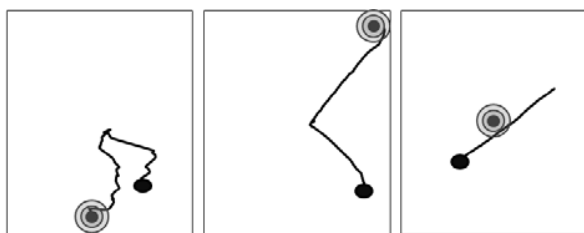


Figura 9 – Exemplo de gestos realizados pelo Participante #3.

#### 4.2.3 Seleccionar alvos na aresta

Uma das supostas barreiras para uma experiência eficaz no uso de ecrãs tácteis é a ausência de elementos físicos a suportar os elementos visuais. É comuns as interfaces móveis apresentarem alvos junto às arestas (ex: fechar janela, atalhos, menu). As arestas do ecrã representam o último suporte táctil (que já não existe em alguns modelos). Queremos perceber se as arestas são um benefício na aquisição de um alvo e qual a melhor opção para os seleccionar.

Do grupo de técnicas apresentado, os utilizadores seleccionaram alvos na aresta através do *Tapping* ou cruzando-os em direcção à borda (*Exiting*). Não foram encontradas diferenças significativas entre estas duas abordagens, sugerindo eficácia semelhante.

As técnicas individuais também não mostraram nenhum efeito da *Posição* (aresta ou não) na eficácia de selecção (*Erros*) (ver secção 4.1). No entanto, um efeito significativo foi encontrado nos *Erros* na comparação entre as abordagens de *Crossing* e *Exiting* nos tamanhos menor ( $F_{1,28}=12.62, p<.01$ ), médio ( $F_{1,28}=12.97, p<.01$ ) e maior ( $F_{1,28}=5.95, p<.05$ ), mostrando que um gesto para um alvo no meio do ecrã é mais eficaz que um gesto em direcção a uma aresta (Figura 4). Isto pode dever-se às restrições impostas no cenário da aresta.

#### 4.2.4 Interagir no centro do ecrã

O centro do ecrã refere-se a todos os alvos que não estejam em contacto com nenhuma aresta. Este grupo representa a maior percentagem da superfície de interacção e vale a pena compreender como podem os utilizadores melhor interagir aí. Neste estudo, os participantes podiam tocar (*Tapping*) ou cruzar (*Crossing*) um alvo no meio do ecrã e ainda realizar gestos direccionais nessa zona. Nesta análise descartámos os gestos diagonais, visto estes já terem sido vistos como catastróficos para esta abordagem. Não encontramos diferenças significativas da *Técnica* nos *Erros*, sugerindo que os utilizadores apresentam eficácia semelhante a interagir no centro do ecrã.

#### 4.2.5 Opiniões dos Utilizadores

Quando questionados acerca da Facilidade de Utilização (numa escala de Likert de 5 pontos), a mediana [quartis] dos valores atribuídos pelos utilizadores foram para o *Exiting* 2[2, 3], *Tapping* 4[4, 4.5], *Crossing* 4[4, 4] e Gestos Direccionais 4[4, 4], mostrando uma ligeira preferência pelo *Tapping*. Esta ideia é reforçada quando os utilizadores foram questionados sobre a sua técnica pre-

ferida (9/15 seleccionaram *Tapping*, 3/15 seleccionaram *Crossing*, e 3/15 seleccionaram *Gestos Direccionais*).

## 5. DISCUSSÃO

Após a análise de cada técnica e da comparação entre elas, estamos agora em posição de responder às questões de investigação propostas no início deste estudo.

### 1. Qual o melhor tamanho para cada técnica?

Para todas as técnicas, os melhores tamanhos foram o médio e o maior (12 e 17 mm) sem diferenças significativas entre ambos. Da perspectiva de alguém que desenha interfaces, alvos com 12 mm serão um bom compromisso devido às restrições de espaço dos dispositivos móveis. Tem de ser feita a excepção à técnica *Exiting*, que apenas se mostrou significativamente diferente entre os tamanhos menor e maior. Assim, fazer um gesto de encontro a um alvo na aresta requer alvos maiores.

Apesar destes resultados, a média de Erros é ainda alta (20%), quando comparada com a reportada para pessoas sem deficiências motoras [6, 11, 12, 13] (menor que 10%). Assim, é necessária mais investigação para melhorar o desempenho de deficientes motores com interfaces em ecrãs tácteis.

### 2. Qual a melhor área do ecrã para cada técnica?

Relativamente aos *Erros*, as áreas do ecrã não afectaram o desempenho dos utilizadores. No entanto, cantos e arestas permitem aos utilizadores seleccionar alvos de forma mais precisa e realizar gestos direccionais mais lineares. Isto sugere que, de facto, as barreiras do ecrã oferecem estabilidade física [Froelich07, Wobbrock03]. Ainda, quando se colocam alvos perto das arestas ou cantos tem de se ter atenção à selecção do tamanho e da técnica, devido às restrições físicas impostas pelo dispositivo. Embora ofereçam estabilidade, podem também restringir o movimento. Os resultados mostraram que cruzar alvos no centro do ecrã é mais fácil que cruzar alvos na aresta.

### 3. Qual a melhor técnica para a população alvo?

Neste estudo, o *Tapping* e o *Crossing* apresentaram-se como as técnicas mais eficazes. No entanto, levando em consideração as opiniões dos utilizadores, o *Tapping* leva ligeira vantagem, visto que 9 em 15 o seleccionaram como método preferido.

### 4. Qual a melhor técnica para cada área?

No centro do ecrã nenhuma técnica se mostrou mais eficaz que as restantes. Isto sugere que os utilizadores atingem eficácia semelhante com qualquer uma delas. No entanto, quando se permite *Gestos Direccionais*, apenas devem ser consideradas 4 direcções (Norte, Sul, Este e Oeste), visto que as diagonais são difíceis de atingir e o elevado número de opções restringe o ângulo atribuído a cada uma. Em relação aos cantos do ecrã, o *Tapping* e *Exiting* têm eficácia semelhante. Já em relação às arestas, o *Tapping* é uma opção mais adequada, visto que realizar gestos nesta área pode ser difícil para pessoas com dificuldades motoras.

### 5. Qual a melhor combinação técnica-tamanho-área?

Levando em consideração as respostas anteriores, e resumizando todas as questões de investigação, podemos concluir que a melhor combinação será o uso de *Tapping*, no centro do ecrã com alvos de tamanho igual ou superior a 12 mm. Esta combinação foi a que mostrou melhores resultados na interação de pessoas tetraplégicas com ecrãs tácteis. No entanto, aplicações específicas podem necessitar de medidas específicas. Quando esta combinação não for praticável, os resultados aqui apresentados dão espaço para alternativas.

### 6. CONCLUSÕES

As interfaces tácteis apresentaram-se como alternativas apelativas às tradicionais interfaces por botões, particularmente para aqueles que não detêm a força ou destreza necessária nos membros superiores [Myers02]. Ainda assim, estes dispositivos impõem novos desafios aos utilizadores, como por exemplo a ausência de estabilidade física, dificultando a selecção eficaz de alvos. Foram desenvolvidas algumas técnicas para melhorar o acesso a interfaces tácteis móveis para pessoas com deficiências motoras [Froehlich07, Wobbrock03]. No entanto, estas foram desenvolvidas sem bases empíricas nem caracterização das capacidades e necessidades dos utilizadores ao interagir com estes dispositivos. É necessário dar um passo atrás e perceber o que conseguem os utilizadores fazer e como.

Os nossos resultados indicam que os utilizadores beneficiam de um melhor conhecimento das suas capacidades e desafios, em particular considerando a técnica de interacção (*Tapping*, *Crossing*, *Exiting* ou *Gestos Direccionais*), tamanho (7, 12, ou 17 mm) e área do ecrã (centro, arestas, cantos, ou proximidade do utilizador).

O *Tapping* mostrou-se como uma técnica promissora quando comparada com métodos alternativos, como o *Crossing* ou os *Gestos Direccionais*. Para além de ser a técnica preferida, os utilizadores foram tão eficazes com esta técnica como com as restantes. Este resultado sugere que uma interface unificada pode ser desenvolvida em benefício de utilizadores com e sem deficiências motoras, com ajustes menores.

Em relação ao tamanho, os 12 mm mostraram-se como um bom diâmetro. No entanto, apenas três tamanhos foram estudados. É importante reconhecer a necessidade de investigação adicional para refinar o conjunto de tamanhos possíveis.

Relativamente a áreas do ecrã, os nossos resultados confirmam que arestas e cantos oferecem estabilidade física. No entanto, não melhoram a eficácia. Por outro lado, existe um efeito da distância do alvo ao utilizador, indicando que os utilizadores são mais eficazes a tocar o ecrã perto do suporte do seu braço. Com técnicas baseadas em gestos, este efeito esbate-se, sugerindo que os utilizadores conseguem adquirir alvos que estejam mais longe, com o apoio da superfície, ajustando o seu movimento.

### 7. AGRADECIMENTOS

Os autores agradecem a todos os participantes deste estudo. O trabalho foi financiado pela FCT através do programa PIDDAC. Tiago Guerreiro e Hugo Nicolau foram apoiados pela FCT, através das bolsas SFRH/BD/28110/2006 e SFRH/BD/46748/2008, respectivamente.

### 8. REFERÊNCIAS

- [Froehlich07] Froehlich, et al. 2007. Barrier pointing: using physical edges to assist target acquisition on mobile device touch screens. In Proc. of ASSETS, 19-26.
- [Guerreiro10] Guerreiro, et al. 2010. Assessing mobile touch interfaces for tetraplegics. In Proc. of MHCI.
- [Lee09] Lee, S., and Zhai, S. 2009. The performance of touch screen soft buttons. In Proc of SIG CHI'09, 309-318.
- [Lin07] Lin, et al. 2007. How do people tap when walking? An empirical investigation of nomadic data entry. In International Journal of Human-Computer Studies. Elsevier, 65, 9, 759-769.
- [Mizobuchi05] Mizobuchi, et al. 2005. Mobile text entry: relationship between walking speed and text input task difficulty. In Proc. of MobileHCI, 122-128.
- [Myers02] Myers, et al. 2002. Using handhelds to help people with motor impairments. In Proc. of ASSETS, 89-96.
- [Parhi06] Parhi, et al. 2006. Target size study for one-handed thumb use on small touchscreen devices. In Proc. of MobileHCI'06, 203-210.
- [Park08] Park, et al. 2008. Touch key design for target selection on a mobile phone. In Proc. of Mobile HCI'08, 423-426.
- [Perry08] Perry, K., and Hourcade, J. 2008. Evaluating one handed thumb tapping on mobile touchscreen devices. In Proc. of Graphics Interface, 57-64.
- [Sears03] Sears, et al. 2003. When computers fade: Pervasive computing and situationally induced impairments and disabilities. In HCI Int'l. Lawrence Erlbaum Associates, 2, 1298-1302.
- [Soukoreff04] Soukoreff, R., and MacKenzie, I. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. In International Journal of Human-Computer Studies. Elsevier, 61, 6, 751-789.
- [Wobbrock08] Wobbrock, J., and Gajos, K.. 2008. Goal crossing with mice and trackballs for people with motor impairments: Performance, submovements, and design directions. In TACCESS, 1, 1, 1-37.
- [Wobbrock03] Wobbrock, et al. 2003. EdgeWrite: a stylus-based text entry method designed for high accuracy and stability of motion. In Proc. of UIST'03, 61-70.



# Interacção do Sénior com Jardim de Lar: Desenvolvimento de um Guião de Entrevista

## Cláudia Mourato Nunes

Centro de Investigação em  
Arquitectura, Urbanismo e  
Design, Faculdade de  
Arquitectura de Lisboa,  
Universidade Técnica de Lisboa  
R. Prof. Cid dos Santos, Pólo  
Universitário, Alto da Ajuda, 1300  
Lisboa  
claudianunes@hotmail.it

## Francisco dos Santos Rebelo

Laboratório de Ergonomia,  
Faculdade de Motricidade  
Humana, Universidade Técnica  
de Lisboa,  
Estrada da Costa, 1499-002 Cruz  
Quebrada, Dafundo  
frebelo@fmh.utl.pt

## Fernando Moreira da Silva

Centro de Investigação em  
Arquitectura, Urbanismo e  
Design, Faculdade de  
Arquitectura de Lisboa,  
Universidade Técnica de Lisboa  
R. Prof. Cid dos Santos, Pólo  
Universitário, Alto da Ajuda,  
1300 Lisboa  
moreiradasilva.fernando@  
gmail.com

## Sumário

*A qualidade dos jardins de lar é um ponto fundamental para a qualidade de vida do idoso. A persecução deste objectivo nem sempre é simples para o projectista, porque existem muitas lacunas na literatura sobre a relação entre as características de um jardim e as motivações, capacidades e limitações do sénior.*

*Neste contexto, este trabalho corresponde a uma fase preliminar de um projecto de investigação mais lato e tem como objectivo a apresentação de uma metodologia para o desenvolvimento de um guião de entrevista (tratando-se de um pré-teste), para se conhecer a forma como os idosos interagem com um jardim de lar, em função das características de jardins existentes. Foram fotografados quatro jardins de lares portugueses e seleccionadas vinte e três fotografias em função dos seguintes critérios: elementos do design e de composição dum jardim, aliados às actividades de lazer. Tendo como ponto de partida estes elementos, desenvolvemos um primeiro guião de uma entrevista com doze perguntas que foram testadas com cinco idosos entrevistados (Estudo Piloto).*

*Durante a entrevista verificámos a necessidade de rever quatro fotografias, em que manifestaram dificuldades em identificar os elementos gráficos; no entanto, para a generalidade dos aspectos, não identificámos problemas de compreensão. Os resultados desta entrevista mostraram que a metodologia utilizada foi eficaz para medir a forma como os idosos manifestam as suas intenções comportamentais no que se refere à sua interacção com os jardins de lar. Após a modificação deste guião, ele será aplicado a uma amostra significativa de idosos, com o objectivo de recolher conhecimento útil na elaboração de cenários para jardins de lar.*

## Palavras-chave

*Design Inclusivo, Design Ergonómico, Sénior, Satisfação e Qualidade de Vida, Jardim de Lar.*

## 1. INTRODUÇÃO

A população idosa portuguesa dedica cerca de duas horas diárias às actividades passivas de descanso ([INE01], [Nunes08]). Os jardins podem ser um local privilegiado para as actividades de lazer e descanso. No entanto, é em actividades de lazer que a população portuguesa regista maior índice de acidentes [Brandão; Brito; Nunes02]. Uma explicação para este problema pode estar relacionado com os problemas associados à senescência do sénior - este experimenta muitas vezes a queda (80% das entradas nos centros de saúde e hospitais portugueses com esta lesão são idosos, devido a contusões/hematomas, com incidência nos membros, e que provoca internamentos prolongados [Brandão; Brito; Nunes02]), perda de equilíbrio, desorientação e confusão, etc. [Burton; Mitchell07] -, daí que a facilidade ou a dificuldade com que realiza as suas actividades no seu quotidiano ser muito influenciada pelas

características do ambiente, ou seja, do design de espaços exteriores [Nunes08].

Não se encontram estudos direccionados para esta questão, que se julga muito pertinente face à actual conjuntura da sociedade ocidental (envelhecimento populacional [UnitedNations09] e institucionalização do idoso [Perista98]) e às lacunas na literatura dos profissionais do «projecto» (arquitectos, arquitectos paisagistas, designers, urbanistas, etc.) sobre a área do design inclusivo e do design ergonómico de espaços exteriores, pelo que se pensa que o presente estudo se venha a revestir num trabalho significativo e inovador. Assim, torna-se importante detectar o real impacte dos espaços verdes (jardins) nas actividades de lazer do idoso utente de lar, de maneira a saber a relação entre as características de um jardim e as capacidades, interesses e limitações do sénior.

A qualidade dos jardins de lar é pois um ponto fundamental para a qualidade de vida do idoso [Nunes02].

Alguns estudos na área da satisfação, bem-estar e qualidade de vida e qualidade dos espaços exteriores foram realizados por especialistas ([Águas;Brandão;Carrelo02], [Alves96], [Alves;Sugiyama06], [Burton;Mitchell07], [Crist99], [Conell04], [Flanagan78], [Flanagan82], [IDGO08], [Maslow70], [Neri07], [Norberg-Schulz86], [Sousa;Galante;Figueiredo03], [Tuan80], [Whitehouse et al.01]), não respondendo, no entanto, directamente a questões ligadas com os jardins de lares, nem na sua maioria expressam a opinião do presente *core* de utente (não abordando o Design Centrado no Utilizador).

Os estudos que se aproximam mais do supramencionado são os estudos publicados por Bite e Lovering, e Barnes e Cooper-Marcus ([Bite;Lovering84] [Barnes;Cooper-Marcus95]), embora não reflectam o panorama nacional (*core* do presente estudo), as pesquisas reportam-se a alguns anos a esta parte, podendo não corresponder à realidade actual, e aponta-se, como crítica principal, o facto dos parâmetros mencionados nos seus estudos serem muito subjectivos, pelo que a presente investigação pretende ir mais longe, quantificando cada variável no campo do Design, tornando-os objectivos e facilmente utilizáveis pelos projectistas.

Refira-se que não é prática comum em *ateliers*, usar-se a metodologia da entrevista mediante questionários de recolha de informação por parte do utente a que se destinam os projectos, não promovendo o processo de interacção «local-utente-actividade» (neste caso, «jardim-senior-lazer»). Os motivos apontados são, entre outros, a falta de conhecimento, sensibilização ou interesse, aliada à vertente económica e/ou falta de recursos humanos especializados (equipa multidisciplinar).

Ao nível do panorama internacional, especialmente no Reino Unido e nos E.U.A. ([Search00], [Anderson90]), existe uma aproximação maior quanto à problemática da concepção de jardins de lar adaptados à terceira idade, comparativamente com o panorama nacional. No entanto, também não têm em conta os princípios do Design Inclusivo (concepção de espaços abrangendo o máximo de utilizadores, independentemente das suas capacidades e atributos bio-psico-sociais, fazendo participar os utentes de uma forma activa no próprio processo projectual), especializando o lar num tipo específico de desvantagem do utente institucionalizado (exemplo geral, e.g., jardim de lar para doentes com Alzheimer [Stevens95] [Cooper-Marcus05]).

Conhecendo as necessidades bio-psico-sociais ([Henriques97/98], [SNR96], [Leuschner05], [Sousa;Galante;Figueiredo03], [Almeida;Pacheco87]) da população idosa portuguesa, e encontrando-se as variáveis/condicionantes de utilização dos espaços, podem-se encontrar soluções ao nível do desenho, de forma a promover o seu bem-estar/satisfação/saúde/qualidade de vida - conceitos considerados indissociáveis ([Nunes08], [Nunes10]).

O estudo apresentado é uma fase preliminar de um projecto de investigação mais lato, correspondendo a uma abordagem metodológica que se consubstancia num pré-teste. O principal objectivo deste estudo é apresentar o desenvolvimento dum guião de entrevista, para apurar os elementos do design, os elementos de composição do jardim e as intenções comportamentais dos idosos relacionadas com actividades de lazer em jardins de lares.

Espera-se que os resultados obtidos possam permitir de uma significativa ajuda às equipas responsáveis por projectos de jardins de lar. Este guião poderá ser usado pelos profissionais que pretendam desenvolver propostas de jardins tendo em conta os princípios do Design Centrado no Utilizador, isto é, colocando a opinião do indivíduo no palco das futuras decisões projectuais, à semelhança do sucedido para o projecto de concepção de jardim público «Jubilee Gardens» (Londres) [Moore05].

## 2. METODOLOGIA

Do ponto de vista metodológico, esta fase do estudo desenvolveu-se em três principais etapas, todas de base qualitativa e não intervencionista (Figura 1):

- Observação directa de jardins de lares;
- Selecção de fotografias de jardins de lares;
- Entrevista com Guião.

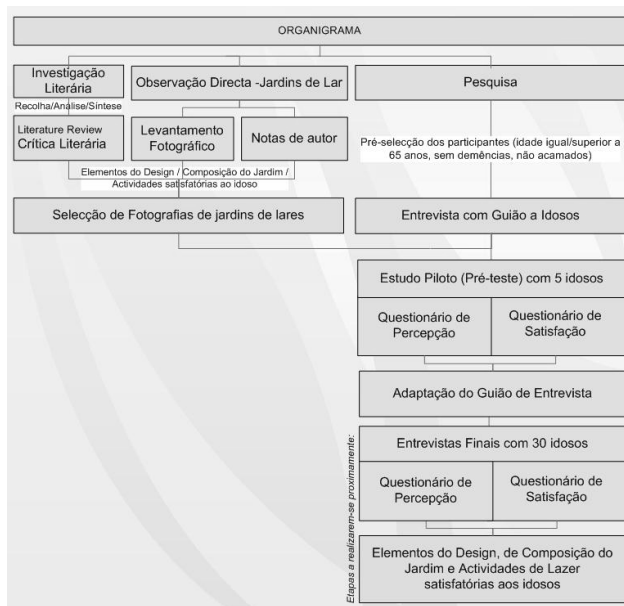


Figura 1: Organograma do processo metodológico.

### 2.1. Observação Directa de Jardins de Lares

Foram observados quatro jardins de lares nos Concelhos da Amadora, Loures, Lourinhã e Palmela. Realizaram-se levantamentos fotográficos (o mais completos possível, fazendo uma cobertura pormenorizada dos espaços ajardinados dos lares) e descreveram-se (notas de autor) os elementos de composição do jardim, os elementos do desenho e os usos/actividades/zonamentos.

### 2.2. Selecção de Fotografias de Jardins de Lares

A escolha das fotografias de autor obedeceu aos seguintes critérios:

- Ter em conta estudos anteriores que classificam a qualidade de vida do idoso no espaço exterior;
- Percepção profissional dos investigadores;
- Seleccionadas *à priori* pelos investigadores, face ao já conhecido através do estado de arte, escolheram-se fotografias de jardins de lar que focam as considerações presentes na Tabela 1 (note-se que numa mesma fotografia várias variáveis e sub-variáveis poderão ser compreendidas).

Variáveis	Sub-variáveis
Características dos Elementos do Desenho	Cor, Textura, Linha, Forma
Princípios do Desenho	Unidade, Variedade, Harmonia, Ênfase, Balanço, Escala, Ritmo
Elementos de Composição do Jardim	Vegetação, Pavimentos/Caminhos, Mobiliário e Equipamento Urbano
Actividades de Lazer	Recreio Activo e Recreio Passivo

**Tabela 1: Elementos de composição (jardim e design) e actividades satisfatórias ao idoso, a apurar no estudo.**

### 2.3. Entrevista com Guião

Usa-se a metodologia de entrevista ao *target/core* do estudo: pessoas institucionalizadas em lares portugueses.

O sénior é pré-seleccionado pelos investigadores e outros profissionais da área dos cuidados de saúde. Isto deve-se ao facto de se pretender idosos com 65 ou mais anos, possuindo pelo menos visão residual (satisfatória para se aferirem os resultados) [Bright;Cook;Harris97], sem demência (que dificultaria o apuramento de dados) e não acamados (impossibilitados de vivenciar o jardim) [Astell&al09].

Através da entrevista, auxiliada pelas fotografias de soluções a jardins de lares, é possível obter indicadores/variáveis de satisfação, tentando saber quais as actividades que os participantes lá têm ou teriam noutros locais (revelando respectivamente, o seu comportamento usual e a sua intenção comportamental). Igualmente, dada a sua frequência a esses espaços, saber o grau de satisfação face à situação apresentada, emitindo comentários vários (tentando-se perceber o porquê da escolha) que conduzirão a pistas para os passos seguintes do estudo.

Em suma, nesta fase importa apurar dados junto dos utilizadores de jardim (cerca de 30 utentes de lar), de modo a que a investigação seja correctamente conduzida (saber as dificuldades sentidas e o design de jardins que preferem). Mas primeiro apurou-se a assertividade do guião junto de cinco idosos institucionalizados em lar (realização de Estudo Piloto no Lar de Idosos da Lourinhã).

A construção do questionário passou por um estudo aquando da análise/crítica literária, conduzindo à elaboração assertiva das questões, com o objectivo de recolher tanto quanto possível, impressões espontâneas e imediatas sobre o objecto do estudo.

Foram colocadas questões o mais compreensíveis possível (como sugerido por uma psicóloga especialista na área da gerontologia, consultada para o efeito), permitindo aos investigadores obterem um *feedback* quanto ao modo como o estudo deveria ser conduzido.

A entrevista teve uma duração aproximada de 40 minutos, e foi realizada em contexto privado e individual (para não expor o idoso a terceiros, e não haver «contaminação» nas respostas por influência doutros).

No início de cada entrevista foi feita a apresentação do investigador, pediu-se o consentimento à sua realização e fez-se um preâmbulo explicativo do assunto a que se destinava (apresentado numa linguagem simples, para fácil entendimento do *core*).

Numa primeira fase, ao entrevistado foram feitas perguntas relativas a imagens de fotografias de jardins de lares (Entrevista com Questionário de Percepção), visualizadas através duma apresentação multimédia, a partir de um computador portátil. Esta técnica permitiu apurar, de uma forma inovadora e atractiva, as respectivas respostas. Também, permitiu ajustar a dimensão das imagens às necessidades de cada observador.

Numa segunda fase da entrevista, foram feitas perguntas abertas e fechadas sobre o respectivo jardim de lar (Entrevista com Questionário de Satisfação).

#### 2.3.1. Protocolo da Entrevista

Segue-se a estrutura do guião usado nas entrevistas aos seniores do Lar de Idosos da Lourinhã.

ASSUNTO: Trabalho universitário sobre o grau de percepção e satisfação sobre os Jardins de Lar. Os resultados obtidos são usados apenas para fins de investigação, sendo processados em conjunto, e ao ser realizada a entrevista é dado consentimento para o efeito. A sua opinião é muito importante, contribuindo de uma forma eficaz para que se venha a usufruir de jardins com qualidade. Note-se que não existem respostas correctas ou incorrectas, apenas se pretende saber a sua opinião acerca das perguntas que se seguem. Agradecemos desde já a sua colaboração.

1. Comente o que vê nas fotografias. [*apresentação multimédia de 23 fotografias de jardins de lar*]
2. O que geralmente faz (actividades) nesse local do jardim do seu lar? [*apresentação multimédia de 6 fotografias tiradas no lar da Lourinhã*]
3. Comente o que faria noutros jardins de lar que se ilustram de seguida. [*apresentação multimédia de 17 fotografias tiradas noutros lares*]
4. Quantas vezes vai ao espaço exterior ajardinado do seu lar?
  - ☐ Mensalmente
  - ☐ Semanalmente
  - ☐ Diariamente, uma vez
  - ☐ Várias vezes ao dia
5. Qual o período do dia em que frequenta o jardim? (*várias hipóteses podem ser escolhidas*)
  - ☐ Manhã
  - ☐ Tarde
  - ☐ Noite

6. Quanto tempo permanece em média no jardim?

- ☐ Menos de 30 minutos
- ☐ 30 min. a 1 hora
- ☐ 1-2 horas
- ☐ 2-4 horas
- ☐ Mais de 4 horas

7. Relativamente às condições meteorológicas, evita ir ao jardim, em dias: (*várias hipóteses podem ser escolhidas*)

- ☐ Chuvosos
- ☐ Frios
- ☐ Ventosos
- ☐ Calor
- ☐ Nublados
- ☐ É indiferente

8. Sente alguma diferença em si após ter permanecido algum tempo no jardim? (*várias hipóteses podem ser escolhidas*)

- ☐ Nenhuma diferença
- ☐ Mais calmo, relaxado, descontraído
- ☐ Mais comunicativo e desinibido
- ☐ Mais alegre, menos deprimido
- ☐ Outras. Quais? \_\_\_\_\_

9. Que características específicas ou qualidades do jardim o fazem sentir bem de modo a frequentá-lo?

10. Globalmente, qual o nível de satisfação ao frequentar o jardim?

- ☐ Muito Satisfatório
- ☐ Satisfatório
- ☐ Indiferente
- ☐ Insatisfatório
- ☐ Muito Insatisfatório

11. Que melhorias introduziria no jardim?

12. Dados Pessoais:

Nome \_\_\_\_\_

Idade \_\_\_\_

Género:

- ☐ Feminino
- ☐ Masculino

Estado civil:

- ☐ Solteiro(a)
- ☐ Casado(a)/junto(a)
- ☐ Viúvo(a)
- ☐ Divorciado(a)/separado(a)

Sector de Actividade Profissional: (*várias hipóteses podem ser escolhidas*)

- ☐ Sector Primário (Agricultura e pescas)
- ☐ Sector Secundário (Operário fabril)
- ☐ Sector Terciário (Empregados de escritório, comércio e serviços)

Deficiência:

- ☐ Não

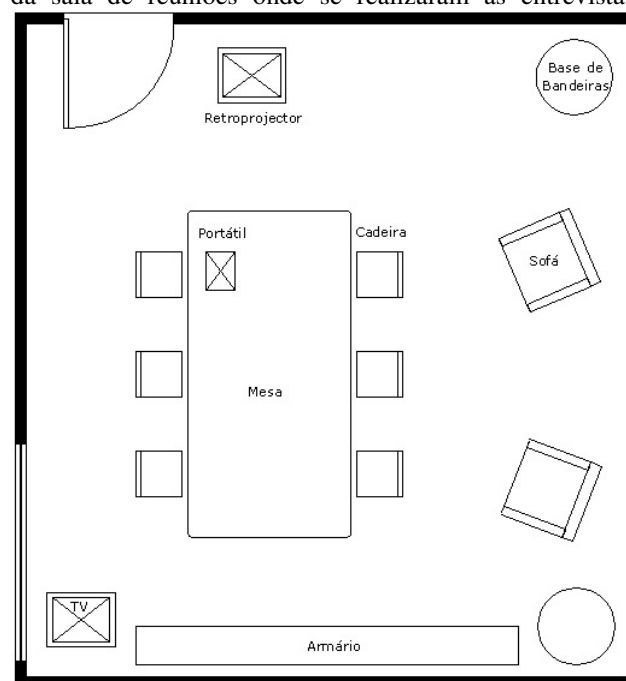
Sim (*permanente ou temporária; várias hipóteses podem ser escolhidas*)

- ☐ Auditiva
- ☐ Visual
- ☐ Física
- ☐ Psíquica

Habilitações Literárias:

- ☐ Instrução Primária
- ☐ Liceu ou Escola Comercial/Industrial
- ☐ Graduação Superior
- ☐ Sem. Mas sabe ler e escrever
- ☐ Sem. Analfabeto

A figura seguinte (Figura 2) ilustra a planta e o inventário da sala de reuniões onde se realizaram as entrevistas.



**Figura 2: Esquema da planta da sala de reuniões (sem escala) onde decorreram as entrevistas.**

A sala (5,5mx6,0m) estava dotada dos meios e condições específicas de funcionamento para o efeito, nomeadamente dispondo de: mesa; cadeiras; computador portátil (elemento externo, trazido pelo entrevistador); TV e mesa de suporte; retroprojector e mesa de suporte; sofás; bases de suporte de bandeiras; armário.

### 2.3.2. Estudo Piloto

O objectivo do Estudo Piloto foi poder testar o guião da entrevista junto de cinco idosos (de um dos quatro lares analisados *à priori*, ou seja, no Lar de Idosos da Lourinhã) para verificar a sua assertividade de modo a apresentá-lo futuramente a 30 participantes.

Todas as fotografias utilizadas foram alvo de apreciação por parte dos cinco idosos (1ª etapa), com o intuito de saber se os elementos de composição lhes eram perceptíveis (através da utilização de um Questionário de

Percepção), de modo a poder-se proceder às etapas seguintes do estudo (vide Tabela 2).

Reconhecimento dos elementos das fotografias pelo participante idoso
+
Confrontação dos resultados com os pré-determinados pelos investigadores
=
Averiguação da interpretação assertiva das fotos pré-seleccionadas (i.e., leitura correcta do envolvimento), com possibilidade de prosseguir com perguntas sobre as mesmas (etapas seguintes)

**Tabela 2: Esquema da apreciação das fotografias.**

Esta fase correspondeu à pergunta 1 do guião de entrevista, apresentado anteriormente.

As fotos seleccionadas procuraram, numa 2ª etapa, saber que actividades os idosos privilegiariam realizar nesses espaços (tentando, por parte dos investigadores, fazer a diferenciação entre actividades de recreio activo e recreio passivo, uma vez saber-se que esse facto conduz a apuramentos diferentes), revelando o comportamento do participante no seu quotidiano e nesse local. Assim, este levantamento de opinião por parte do sénior pretendeu, a partir dum caso concreto, apurar e validar as actividades mais referidas como as idealmente realizadas por si, em jardins de lares portugueses. Esta fase correspondeu à pergunta 2.

A 3ª etapa correspondeu à análise de espaços (fotos de espaços exteriores ajardinados de várias instituições/lares), do que fariam nesses locais (jardins desconhecidos, por retratarem outros lares), revelando uma intenção comportamental, correspondendo à pergunta 3.

Pretendeu-se de seguida saber o grau de satisfação (mediante a utilização de um Questionário de Satisfação) dos idosos sobre os espaços verdes em que são utentes e igualmente aferir os seus dados pessoais. Esta 4ª fase correspondeu às perguntas 4 a 12.

### 2.3.3. Adaptação do Guião de Entrevista

#### Aferição dos Resultados:

O estudo piloto teve o seguinte tratamento de dados:

- Análise das respostas das entrevistas;
- Criação de legendas, de acordo com a variável em análise (Característica dos Elementos do Desenho, Princípios do Desenho, Elementos de Composição de um jardim, Actividades de Lazer realizadas num jardim);
- Identificação das variáveis (e sub-variáveis) mencionadas pelo participante e atribuição do respectivo valor da legenda criada à partida (acrescentar de sub-variáveis à medida que se interpretavam os dados, reformulando a legenda inicial apresentada na Tabela 1);
- Determinação da moda para cada uma das sub-variáveis presentes nas fotografias (perguntas 1 a 3). Visto as variáveis do estudo serem qualitativas, usa-se a moda por esta representar o(s) valor(es) que regista(m) o maior número de observações (valores mais frequentes);

- Determinação da moda das sub-variáveis mencionadas pelo participante, a cada pergunta efectuada (perguntas 4 a 12). Determinação da média, no caso das idades dos entrevistados;
- Conclusões do Estudo Piloto: teste da assertividade do guião junto de cinco idosos, a apresentar futuramente a 30 participantes.

#### Resultados observados face ao objectivo do Estudo Piloto:

- Foram identificadas com sucesso as características dos elementos do desenho mais focadas pelo idoso. Igualmente, os elementos de composição do jardim e as actividades de lazer;
- Não se conseguiram apurar os princípios do desenho (apenas algumas respostas se reportaram ao elemento de escala, presente nos elementos do jardim, sendo previsível apenas apurar estes parâmetros, aquando da fase de estudo com recurso à utilização da ferramenta da Realidade Virtual);
- Quatro fotografias, das vinte e três apresentadas, suscitaram dúvidas em alguns idosos, tendo sido necessário adaptar o guião:
  - Depósito de gás (objecto não familiar a alguns idosos, e imagem cujo ponto fulcral de interpretação pretendia ser o do cultivo de plantas e a vegetação «amenizadora» de elementos construídos);
  - Equipamento de barras (objecto não familiar no quotidiano do sénior do lar em questão, interpretado, por exemplo, como latada por um dos idosos participantes – objecto que lhe é mais familiar);
  - Zona de Sol, seguida de sombra (interpretada como água ou terreno, o que pode remeter para um estado de envelhecimento a nível da visão do idoso inquirido, e que corrobora o conhecido através do estado de arte, i.e., incapacidade de percepção de profundidades de sombras contrastadas com luz [AIA85;Brawley01;Campbell05;Carsten85;Harrington 93]);
  - Elemento de água (compreendido como pavimento, o que pode igualmente remeter para um estado de envelhecimento a nível da visão).

Foram alteradas as quatro imagens que revelaram problemas de compreensão por parte de alguns participantes. Dado ter-se efectuado um levantamento fotográfico dos jardins de lar observados, existiam outras fotografias que estavam em condições de substituir as anteriores (sendo necessário testá-las primeiramente para as validar). Assim, realizaram-se novas entrevistas aos cinco idosos, de modo a aferir a leitura correcta dos envoltórios. Os resultados, face às novas fotografias apresentadas, foram muito satisfatórios no caso do elemento de água (compreendido como tal) e do depósito (visto este último não ter sido objecto de crítica que pusesse em causa a validação de toda a imagem apresentada). Quanto à fotografia que apresentava zona de Sol, perto de zona fronteira com a de sombra, um dos entrevistados (dado o problema de visão) interpretou a

sombra como modificação de nível no pavimento (julgando serem raízes). A fotografia do equipamento de barras foi substituída por uma de argolas – o resultado foi o mesmo, pelo que se conclui que não se trata de todo de um objecto familiar aos entrevistados; no entanto, não significa que não seja interpretado por outro idoso que lhe atribua o significado correcto (nomeadamente se o tiver experimentado no seu percurso de vida).

Em suma, os resultados das entrevistas mostraram que a metodologia utilizada foi eficaz para medir a forma como os idosos manifestam as suas intenções comportamentais e comportamentos usuais, para interagir com os jardins de lar.

Tal como referido, o trabalho apresentado encontra-se numa fase preliminar. Assim, e embora os resultados obtidos não sejam significativos, nem que tal seja relevante nesta fase de estudo, deixamos algumas considerações que ilustram a opinião dos seniores entrevistados. Segue-se a nova legenda (em alteração à Tabela 1) face às respostas referidas pelos idosos (Tabela 3):

Elementos do Desenho	Princípios do Desenho	Elementos de Composição do Jardim	Actividades de lazer
a-cor b- textura (material) c-linha d-forma	A-unidade B-variedade C-harmonia D-ênfase E-balanço F-escala G-ritmo	1-vegetação 2-pavimento 3-percurso (caminho, escada) 4-mobiliário e equipamento urbano 5-construído/edificado 6-paisagem 7-vedação/muro/cerca 8-elemento de água 9-condições bioclimáticas vs. iluminação natural 10-outros	i- recreio activo: 1-brincar 2-jogar 3-jardinagem (e afins) 4-bricolage 5-exercício físico 6-outros ii- recreio passivo: 1-contemplar 2-sentar/estadia 3-passear 4-comer 5-conviver/conversar 6-ler/ouvir música/cantar 7-outros

**Tabela 3: Nova legenda, sobre os elementos de composição (jardim, design) e actividades satisfatórias ao idoso, apurada no presente estudo.**

Os resultados face à **pergunta 1** (*Observação de fotografias para identificação dos elementos do desenho e dos princípios do desenho, e de composição dum jardim*), do guião de entrevista foram:

O elemento do desenho mais referido é a cor, associado aos elementos de composição de jardim: Vegetação, seguindo-se construído/edificado e por fim mobiliário e equipamento urbano.

Os elementos de composição de jardim mais observados pelos participantes foram (por ordem decrescente de importância):

- Vegetação;
- Construído/edificado;
- Mobiliário e equipamento urbano;
- Percurso;
- Vedação/muro/cerca e elemento de água.

**Pergunta 2:** *Observação de fotografias (actividades realizadas no lar onde vivem – comportamento usual):*

Os idosos realizam jardinagem, passeiam e desenvolvem outras actividades de lazer no espaço exterior ao seu lar.

**Pergunta 3:** *Observação de fotografias (possíveis actividades realizadas no jardim de lar apresentado – intenção comportamental):*

Actividades mais realizadas: jardinagem, segue-se o passear e o sentar/estadia, passear (passagem), contemplar vistas e outras actividades, finalmente conviver/conversar e brincar, por exemplo, com crianças.

Resultados das **perguntas 4-11**:

- Os idosos inquiridos frequentam o seu jardim entre uma a várias vezes ao dia;
- É mais usual frequentarem o jardim no período da tarde;
- Mais frequente a estadia de 1-2 horas no jardim;
- Dias frios, demasiado quentes e chuvosos inibem a visita ao jardim;
- Após frequentarem o jardim do lar não ficam indiferentes: sentem-se mais alegres/menos deprimidos e mais calmos/relaxados/ descontraídos;
- A existência de vegetação e de equipamentos/mobiliário urbano no jardim é um factor importante para a sua satisfação;
- Quanto ao grau de satisfação face ao jardim de lar, os resultados são muito díspares: de muito satisfeito a muito insatisfeito;
- A referência à vegetação mostrou ser o elemento mais referido pelos idosos de modo a melhorar o seu jardim.

Da **pergunta 12**, do guião de entrevista, aferiram-se os seguintes dados: A média de idades dos participantes é 85 anos (idade mínima:81; idade máxima:93), entrevistas conduzidas a 3 idosos e 2 idosas, viúvos/as, que tiveram várias actividades ao longo do seu percurso profissional, mas tendo em comum o sector agrícola, sem habilitações literárias ou muito baixas (instrução primária), e contando com algum histórico a nível de deficiências (mas não muito acentuadas, e como tal permitindo o apuramento dos dados) principalmente do foro biológico (visão, audição e mobilidade).

### 3. CONCLUSÃO

O principal problema que actualmente existe e que tem consequência para o *design* de jardins, é conhecer de uma forma objectiva as necessidades, dificuldades e motivações dos transeuntes, ao qual o presente estudo se pretende direccionar para a problemática do idoso utente de lar.

Assim, torna-se necessário incluir a satisfação (e qualidade de vida) do sénior como um dos pressupostos para se concretizarem espaços exteriores inclusivos (assegurando a utilização plena de jardins, independentemente das capacidades dos transeuntes) e ergonómicos (em que se desenvolvem interacções positivas entre os vectores



«jardim-idoso-actividade», em condições de segurança, conforto e eficiência do sistema).

Como referido anteriormente, os resultados desta entrevista com guião mostraram que a metodologia utilizada foi eficaz para medir a forma como os idosos manifestam as suas intenções comportamentais para interagir com os jardins de lar. No entanto, apesar de em estudos gerais ser possível obterem-se dados relativos aos princípios universais do desenho, em casos específicos, como os do estudo apresentado, não o é. Outro estudo passará por usar a ferramenta da Realidade Virtual para os apurar em cenários a serem desenvolvidos em três dimensões, colocando os participantes a interagir com os jardins de lar.

Quanto ao apuramento das actividades que os idosos mencionavam realizar face à apresentação das fotografias, estes referiam por vezes usar o espaço individualmente e/ou em grupo, mas que não foram diferenciadas na Tabela 3 e no tratamento estatístico.

O passo seguinte conduzirá à implementação das entrevistas junto de 30 idosos residentes em lar, tendo em conta os seguintes objectivos:

- Analisar as respostas das entrevistas;
- Criar a legenda, de acordo com a variável em análise;
- Identificar as variáveis (e sub-variáveis) mencionadas pelo participante e realizar o devido tratamento estatístico, a fim de se obterem os elementos do design e de composição do jardim, considerados mais significativos/satisfatórios no decurso das actividades dos idosos (para construção de cenários em Realidade Virtual).

No apuramento das actividades que os idosos privilegiam no seu dia-a-dia, tentar-se-á realizar um tratamento estatístico diferenciado quanto às actividades que realizam no espaço individualmente e em grupo.

#### 4. AGRADECIMENTOS

Este trabalho foi financiado por uma Bolsa de Doutoramento atribuída pela FCT - Fundação para a Ciência e a Tecnologia SFRH/BD/61313/2009.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

[Águas;Brandão;Carrelo02] ÁGUAS, S.; BRANDÃO, P.; CARRELO, M. (2002). *O Chão da Cidade, Guia de Avaliação do Design de Espaço Público*. Lisboa: Centro Português de Design, pp. 186-193.

[AIA85;Brawley01;Campbell05;Carsten85;Harrington93] AMERICAN INSTITUTE OF ARCHITECTS (1985). *Design For Aging: An Architect's Guide*. Washington, DC: AIA Press; BRAWLEY, E. (2001). Environmental Design For Alzheimer's Disease: A Quality Of Life Issue. *Aging and Mental Health*, 5 (Suppl 1), S79-S83; CAMPBELL, S. (2005). *Deteriorating Vision, Falls and Older People: The Links*. Glasgow: Visibility; CARSTEN, D. (1985). *Site Planning and Design for the Elderly: Issues, Guidelines and Alternatives*. New York: Van Nostrand Reinhold; HARRINGTON, T. (1993). *Perception and Physiological Physiology in Designing*

*For Older People With Cognitive and Affective Disorders*. Eindhoven: Institute For Gerontology, University of Eindhoven (cited in BURTON, E.; MITCHELL, L. *Inclusive Urban Design-Streets for Life*. Oxford: Elsevier, Architectural Press, 2007, pp. 25-26).

[Almeida;Pacheco87] ALMEIDA, N.; PACHECO, D. (1987). *Anais Portugueses de Saúde Mental - Assistência Psiquiátrica aos Lares de Idosos do Caramulo*. Lisboa: Edição da Direcção da Sociedade Portuguesa para o Estudo da Saúde Mental, Ano 3, Vol. 3, p. 157.

[Alves96] ALVES, F.M.B (1996). *Avaliação da Qualidade do Espaço Público Urbano. Proposta Metodológica*. Lisboa: Instituto Superior Técnico, Universidade Técnica de Lisboa, pp. 290-300.

[Alves;Sugiyama06] ALVES, S.; SUGIYAMA, T. (2006, Dec.). Literature Review on the Benefits of Access to Outdoor Environments for Older People. *Openspace Research Centre*, Edinburgh College of Art. Acedido em 2010 (in [http://www.idgo.ac.uk/useful\\_resources/for\\_other\\_researchers.htm#contact](http://www.idgo.ac.uk/useful_resources/for_other_researchers.htm#contact)).

[Anderson90] ANDERSON, T. (1990, April). Provision For Elderly Mentally Ill. Environment and Occupational Therapy, *Landscape Design*, pp. 22-24.

[Astell et al.09] ASTELL, A.; ALM, N.; DYE, R.; ELLIS, M.; GOWANS, G.; VAUGHAN, P. (2009, April). Involving older people with dementia and their carers in designing computer based support systems: some methodological considerations. *Journal Universal Access in the Information Society*. Vol. 8, n.1. Springer Berlin/Heidelberg, pp. 49-58.

[Barnes;Cooper-Marcus95] BARNES, M.I.; COOPER-MARCUS, C. (1995). *Gardens in Healthcare Facilities: Uses, Therapeutic Benefits and Design Recommendation*. Berkeley: University of California, pp.6-9, 67, 68.

[Bite;Lovering84] BITE, I.; LOVERING, M. J. (1984, Nov.-Dec.). Design Open Doors for the Elderly. *Landscape Architecture*, pp. 79-81.

[Brandão;Brito;Nunes02] BRANDÃO, J.; BRITO, D.; NUNES, B. (2002). *Sistema Adelia 2002-Acidentes domésticos e de Lazer: Informação Adequada-Relatório Final*. Lisboa: ONSA, Instituto Nacional de Saúde Dr. Ricardo Jorge, Ministério da Saúde, pp.14, 19-24, 26. Acedido em 2010 (in [https://webgate.ec.europa.eu/idb/documents/VI\\_2\\_4\\_National\\_Report\\_Portugal.pdf](https://webgate.ec.europa.eu/idb/documents/VI_2_4_National_Report_Portugal.pdf)).

[Bright;Cook;Harris97] BRIGHT, K.T.; COOK, G.; HARRIS, J. (1997). Colour Selection and the Visually Impaired - A Design Guide for Building Refurbishment. Research Group for Inclusive Environments. University of Reading. Acedido em 2009 (in <http://www.rdg.ac.uk> e <http://www.extra.rdg.ac.uk/>).

[Burton;Mitchell07] BURTON, E.; MITCHELL, L. (2007) *Inclusive Urban Design-Streets for Life*. Oxford: Elsevier, Architectural Press, pp.33-165.

[Conell04] CONNELL, J. (2004). The purest of human pleasures: the characteristics and motivations of garden

- visitors in Great Britain. Elsevier *Tourism Management* 25, pp.229–247.
- [Cooper-Marcus05] COOPER-MARCUS, C. (2005, March). No Ordinary Garden. *Landscape Architecture*, Vol. 95, Nº 3, pp. 26-39.
- [Crist99] CRIST, P.A. “Does Quality of Life Vary with Different Types of Housing Among Older Persons? A Pilot Study” in CARLSON, L.; TAIRA, E.D. [Eds] (1999). *Aging in Place: Designing, Adapting, and Enhancing the Home Environment*. New York: The Haworth Press, p.102.
- [Flanagan78] FLANAGAN, J. C. (1978). A research approach to improving our quality of life. *American Psychologist*; 33: 138–147 (cited in CRIST, P.A. “Does Quality of Life Vary with Different Types of Housing Among Older Persons? A Pilot Study” in CARLSON, L.; TAIRA, E.D. [Eds] (1999). *Aging in Place: Designing, Adapting, and Enhancing the Home Environment*. New York: The Haworth Press, p.103).
- [Flanagan82] FLANAGAN, J. C. (1982). Measurement of quality of life: Current state of the art. *Archives of Physical and Medical Rehabilitation* 1982; 63: 56–59 (cited in CRIST, P.A. “Does Quality of Life Vary with Different Types of Housing Among Older Persons? A Pilot Study” in CARLSON, L.; TAIRA, E.D. [Eds] (1999). *Aging in Place: Designing, Adapting, and Enhancing the Home Environment*. New York: The Haworth Press, p.103).
- [Henriques97/98] HENRIQUES, S. (1997/8). *Idosos, Satisfação de Vida e Actividade (Monólogo Clínico)*. Lisboa: Instituto Superior de Psicologia Aplicada, p. 157.
- [IDGO08] INCLUSIVE DESIGN FOR GETTING OUTDOORS. Acedido em 2010 (in [http://www.idgo.ac.uk/older\\_people\\_outdoors/index.htm](http://www.idgo.ac.uk/older_people_outdoors/index.htm)).
- [INE01] INE (2001). *Inquérito à Ocupação do Tempo*. Lisboa: Instituto Nacional de Estatísticas, p. 24.
- [Leuschner05] LEUSCHNER, A. (3-5Nov. 2005). Recomendações do Grupo de Trabalho sobre Saúde Mental no Envelhecimento e Pessoas Idosas. *A cidade e os mais velhos - mais saúde, mais integração e melhor integração (3º Congresso)*. Torre do Tombo: Associação Portuguesa de Psicogerontologia, p.5.
- [Maslow70] MASLOW, A. H. (1970). *Motivation and Personality*. New York: Harper & Row (cited in NUNES, C. “O Idoso, O Lugar e a Satisfação” in TRIGUEIROS, C. [Coord.] (2010). *Uma Utopia Sustentável: Arquitectura e Urbanismo no Espaço Lusófono: que Futuro?* Lisboa: Faculdade de Arquitectura da Universidade Técnica de Lisboa, p.625; e cited in NUNES, C. “Vegetação vs. Satisfação” in *Revista Convergências* nº5, 01-06-2010, Castelo Branco: Escola Superior de Artes Aplicadas in <http://convergencias.esart.ipcb.pt/artigo/77>).
- [Moore05] MOORE, R. (2005). *Jubilee Gardens – Join the celebration and help Jubilee Gardens Blossom, 9th May 2005*, Community Celebration. Jubilee Gardens Steering Group. Acedido em 2008 (in <http://www.jubileegardens.org.uk>).
- [Neri07] NERI, A.L. (2007). *Desenvolvimento e Envelhecimento. Perspectivas biológicas, psicológicas e sociológicas*. Coleção Vivacidade. São Paulo: Papirus, pp. 164-172.
- [Norberg-Schulz86] NORBERG-SCHULZ, C. (1986). *Genius Loci – Paesaggio Ambiente Architettura*. Milano: Electa, p. 202.
- [Nunes02] NUNES, C. (2002). *O Significado das Plantas – Projecto de Concepção de Jardim Terapêutico para Lar de Idosos*. Lisboa: Instituto Superior de Agronomia, Universidade Técnica de Lisboa, pp.110-116.
- [Nunes08] NUNES, C (2008). *Design de Espaços Exteriores. Concepção de Parques e Jardins Adaptados à Terceira Idade*. Lisboa: Faculdade de Arquitectura, Universidade Técnica de Lisboa, pp. 232, 288-293.
- [Nunes10] NUNES, C. “O Idoso, O Lugar e a Satisfação” in TRIGUEIROS, C. [Coord.] (2010). *Uma Utopia Sustentável: Arquitectura e Urbanismo no Espaço Lusófono: que Futuro?* Lisboa: Faculdade de Arquitectura da Universidade Técnica de Lisboa, pp.625-626.
- [Perista98] PERISTA, H. [Coord.] (1998). *Levantamento das Necessidades Sociais das Pessoas Idosas em Contexto Local*. Lisboa: Direcção Geral da Acção Social – Núcleo de Documentação Técnica e Divulgação, pp.54-56.
- [Search00] SEARCH, G. (2000). *The Small Garden Handbook*. England: Island Books, pp. 88-91.
- [SNR96] SNR (1996). *Inquéritos Nacionais às Incapacidades, Deficiências e Desvantagens - Resultados Globais - Cadernos SNR Nº9*. Lisboa: Secretariado Nacional da Reabilitação, p. 133.
- [Sousa;Galante;Figueiredo03] SOUSA, L; GALANTE, H.; FIGUEIREDO, D. (2003). Qualidade de Vida e Bem-Estar dos Idosos: Um Estudo Exploratório na População Portuguesa. *Saúde Pública*:37(3):364-71, p.368.
- [Stevens95] STEVENS, M. (1995, Jan.). *Landscape Architecture*, Vol.85, nº 1, pp. 77-79.
- [Tuan80] YI-FU TUAN (1980). *Topofilia: um estudo da percepção, atitudes e valores do meio ambiente*. São Paulo: Ed. Difel, p. 107.
- [UnitedNations09] UNITED NATIONS. *World Population Ageing 2009*. New York: Department of Economic and Social Affairs/Population Division, 2009, p.xi. Acedido em 2010 (in [http://www.un.org/esa/population/publications/WPA2009/WPA2009\\_WorkingPaper.pdf](http://www.un.org/esa/population/publications/WPA2009/WPA2009_WorkingPaper.pdf)).
- [Whitehouse et al.01] WHITEHOUSE, A.; VARNI, J.W.; SEID, M.; COOPER-MARCUS, C.; ENSBERG, M.J.; JACOBS, J.R.; MEHLENBECK, R.S. (2001). Evaluation a Children’s Hospital Garden Environment: Utilization and Consumer Satisfaction. *Journal of Environmental Psychology* 21, pp. 301-314.

# Reabilitar o Processo de Reabilitação

Teresa Gama Tiago Guerreiro Hugo Nicolau Joaquim Jorge  
 Instituto Superior Técnico / INESC-ID  
 Rua Alves Redol, 9, 100-029, Lisboa  
 mariateresagama@gmail.com, {tjvg,hman,jaj}@vimmi.inesc-id.pt

## Sumário

*O processo de reabilitação actual caracteriza-se pela sua longa duração e natureza desmotivante. Porém, é uma actividade indispensável para a recuperação de pacientes tetraplégicos. O objectivo deste trabalho é tornar a fisioterapia num processo mais divertido e aliciante para os utilizadores. A primeira contribuição deste artigo consiste numa descrição detalhada do processo tradicional de fisioterapia, nomeadamente na caracterização e compreensão dos exercícios mais relevantes. Em segundo lugar, e tendo em conta as necessidades dos utilizadores, retiramos algumas implicações para o desenho de plataformas tecnológicas. Em seguida, apresentamos a nossa abordagem, que conjuga elementos e processos do mundo real com elementos virtuais, podendo assim oferecer aos utilizadores uma experiência mais rica e envolvente, e propomos um conjunto de soluções tecnológicas que poderão tornar a fisioterapia numa actividade mais divertida.*

## Palavras-chave

*Reabilitação, Fisioterapia, Ambiente Virtual, Jogos, Recuperação Motora*

## 1. INTRODUÇÃO

A medula espinal é o nervo mais longo do corpo humano e é responsável pelo controlo das actividades involuntárias do organismo: ritmo cardíaco, respiração ou actividade muscular. Assim, as lesões na medula espinal são sempre muito sérias, visto poderem alterar o sistema cárdio-respiratório ou levar à perda dos movimentos dos membros do corpo humano [Merino08].

A fisioterapia surge, assim, como um dos processos terapêuticos mais relevantes para a recuperação dos pacientes que sofrem uma lesão na medula espinal. O objectivo da fisioterapia é devolver alguma qualidade de vida aos pacientes através do treino de movimentos e reacções, de forma a assegurar uma maior independência e controlo do sistema respiratório [Andrade03].

No entanto, a fisioterapia acaba por ser sempre um processo longo, árduo e aborrecido, sendo que a maioria dos pacientes paraplégicos e tetraplégicos tem de repetir diariamente os mesmos exercícios de forma a assegurar o seu desenvolvimento. Assim, encontrar novas e apelativas formas de fazer fisioterapia é uma tarefa muito importante para facilitar a rotina diária dos pacientes e aumentar a sua motivação e envolvimento neste processo [Rand04].

Este trabalho vem ao encontro deste objectivo. Através da utilização de novas tecnologias, em particular recorrendo ao uso de jogos e ambientes virtuais, pretende-se encontrar novas formas de fazer fisioterapia. A utilização destas soluções tem, para além do aspecto de entretenimento, algumas vantagens associadas: permitir uma aprendizagem prática e activa; possibilidade de medir objectivamente o desempenho dos utilizadores; incitar à

competição e à socialização; facilidade de adaptar as características do ambiente virtual às capacidades do utilizador. Contudo, para atingir o nosso objectivo é necessária uma análise profunda do processo de reabilitação actual, para que, ao contrário de outros sistemas [Kizony06, Foulds07, Sisto08], seja possível criar novos métodos de terapia através de um desenho informado.

Neste artigo será caracterizada a fisioterapia actual, e descritas as implicações para a criação de novas soluções tecnológicas. Posteriormente, serão mostradas sugestões de desenho e cenários de interacção que permitam tornar este processo mais divertido e motivante.

## 2. UM OLHAR SOBRE A REABILITAÇÃO

De forma a compreender as necessidades actuais da fisioterapia e de todo o processo de reabilitação, foi efectuado um estudo no Lar da Cruz Vermelha, Quinta das Conchas. Os principais objectivos eram: identificar os potenciais utilizadores de um novo sistema terapia; perceber



**Figura 1 – Paciente a executar o exercício de rodar a bola pelas pernas.**



como é feita a fisioterapia, hoje em dia; e identificar as principais dificuldades neste processo.

Relativamente aos utilizadores, a amostra obtida foi de 7 pacientes tetraplégicos. Todos tinham entre 18 e 41 anos, com lesões entre os níveis C4 e C7, resultando num controlo residual dos seus membros superiores.

Em relação ao uso de tecnologia, apenas um dos pacientes não tinha computador. Todos os restantes usavam este tipo de aparelhos com frequência. Quanto a jogos, os preferidos e passíveis de serem jogados, são jogos de paciência (e.g. *tetris*, *puzzle bubble*, *majong*), acção (e.g. *counter strike*), jogos de carros e, mais recentemente, jogos associados a redes sociais (e.g. *farmville* da plataforma *facebook*).

Relativamente ao processo de fisioterapia, estávamos interessados em perceber e caracterizar os exercícios realizados, nomeadamente o seu objectivo, materiais utilizados e satisfação dos pacientes. Ao observarmos as dificuldades e motivação dos utilizadores na realização dos exercícios, ficou claro que existe uma relação directa entre o seu desempenho e satisfação, sendo que as maiores dificuldades surgiram nos exercícios que os pacientes menos gostavam.

Nas secções seguintes serão apresentados os exercícios mais relevantes neste processo, tendo em conta o material usado.

### 2.1 Exercícios com a bola de ginástica

Por se assemelharem a um jogo, este tipo de exercícios são os preferidos dos pacientes. Caracterizam-se por movimentos curtos, trabalhando essencialmente o controlo e precisão dos membros superiores e tronco.

#### *Rodar bola sentado*

Sentado, com a terapeuta à sua frente, inclinando-se para o lado e apoiando-se no braço, com a outra mão fazer a bola rodar nas quatro principais direcções.

#### *Bola pelas pernas*

Sentado, mudar a bola de um lado para o outro, rodando-a pelas pernas (Figura 1).

#### *Passe de peito*

Sentado, passar a bola horizontalmente à terapeuta.



Figura 2 – Paciente a executar o exercício Volei-fit-ball.

#### *Três bolas*

Sucessivamente, passar e receber três bolas em passe de peito ou em passe em arco para a terapeuta.

#### *Fitball de apoio*

Sentado, encostado a uma fitball meio vazia, passar em passe de peito a bola para a terapeuta.

#### *Rodar bola deitado*

Deitado lateralmente na cama, com uma mão por cima da bola, fazê-la rodar em todas as direcções.

### 2.2 Exercícios com a *squeezeball*

A *squeezeball* tem as mesmas vantagens que a bola de ginástica, no entanto é muito menos versátil no que respeita à capacidade de interacção do paciente, visto haver apenas um pequeno número de pacientes capazes de segurar um objecto tão pequeno como este., com as mãos. De facto, durante este estudo foi apenas observado um exercício que recorria à *squeezeball*.

#### *Squeeze deitado*

Deitado na cama e apoiado apenas num cotovelo, passar a *squeezeball* para a terapeuta, sentada a 50 cm de distância, empurrando a bola com a mão disponível.

### 2.3 Exercícios com a *fitball*

Ainda na secção dos exercícios que parecem um jogo, outro aparelho muito utilizado é a *fitball*. Esta enorme bola, cheia de ar, não permite movimentos de preensão, no entanto, treina os movimentos largos que são mais facilmente alcançáveis pelos pacientes. Para mais, alguns dos exercícios treinam muito o equilíbrio, que é uma característica em falta nos pacientes tetraplégicos.

#### *Gatas para joelhos*

Com o auxílio da *fitball* como apoio dianteiro, passar da posição de gatas para a posição de joelhos.

#### *Meia vazia*

Sentado, apoiado dorsalmente na *fitball* meia vazia, fazer força a partir dos músculos abdominais para passar à posição tradicional de sentado, e posteriormente regressar à posição inicial.

#### *Sentar*

Aguentar-se sentado na *fitball*, sem se desequilibrar.

#### *Mãos na fitball*

Sentado na cama, com a *fitball* no chão e entre as pernas, passar ambas as mãos em simultâneo da cama para a *fitball* e seguidamente regressar à posição inicial.

#### *Volei-fit-ball*

Sentado na cama e utilizando a *fitball*, jogar voleibol com a terapeuta, que está em pé à sua frente (Figura 2).

### 2.4 Exercícios nas paralelas

As paralelas são o aparelho no qual os pacientes sentem mais medo, derivado do equilíbrio exigido. Tal como ocorre com a *squeezeball*, a variedade de exercícios que podem ser realizados é muito limitada. Durante o período de estudo, apenas observámos um exercício.

### *De pé*

Com a terapeuta sentada à sua frente aguentar-se de pé, mantendo as mãos apoiadas nas paralelas e, sucessivamente, passar cada uma das mãos para os ombros da terapeuta.

### **2.5 Exercícios na cadeira de rodas**

Os exercícios na cadeira de rodas, apesar de terem uma baixa variedade, são os preferidos dos pacientes, devido ao seu elevado desempenho e pelo facto de sentirem alguma liberdade por se poderem deslocar.

#### *Gincana*

Sentado na cadeira de rodas efectuar um caminho previamente especificado pelas terapeutas

### **2.6 Exercícios na cama**

Os exercícios na cama são caracterizados pela baixa aceitação por parte dos pacientes, derivado do facto de este aparelho ser o que os magoa mais, como consequência de terem uma grande parte do corpo em contacto com a cama. No entanto, são muito úteis, visto poderem-lhes fornecer um grande nível de independência, à noite e de manhã, quando se estão a deitar e a levantar das suas próprias camas.

#### *Enrolamento tradicional*

Deitado na cama, rodar o corpo para ambos os lados.

#### *Toque na cama*

Sentado na cama, com a terapeuta sentada em frente, com as mãos apoiadas na cama, levantar cada uma das mãos sucessivamente, mantendo o equilíbrio e a postura.

#### *Inclinar de tronco*

Sentado na cama, com o as pernas esticadas, inclinar sucessivamente o tronco para a frente e para trás.

#### *Braço fixo*

Sentado na cama, esticar cada braço, para a frente e para trás, mantendo o ombro fixo.

#### *Braços esticados*

Levantar ambos os braços até à altura dos ombros, mantendo-os esticados para a frente.

#### *Mãos andantes*

Sentado, apoiado com as mãos atrás, andar com elas para a frente, mantendo o tronco direito e sem se desequilibrar ou dobrar os braços.

#### *Rodar de braços lateralmente*

Deitado de barriga para cima, com pesos de pulso, rodar os braços lateralmente, mantendo-os esticados, desde as pernas até cima da cabeça.

#### *Rodar de braços por cima*

Deitado de barriga para cima, com pesos de pulso, rodar os braços por cima do corpo, desde as pernas até ao cimo da cabeça, mantendo os braços sempre esticados.

#### *Enrolamento de pernas cruzadas*

Deitado na cama, rodar todo o corpo com as pernas cruzadas e levantar-se para a posição sentada.

### *Escorregar*

Deitado de barriga para baixo, com a ajuda dos braços, escorregar pela cama.

### *Articulações*

Sentado com a terapeuta em frente a uma distância de aproximadamente 50 cm, tentar tocar nas suas mãos, ombros, cotovelos ou tornozelos.

## **3. FUNREHAB**

Após termos uma caracterização detalhada dos exercícios de fisioterapia, e com o objectivo de os tornar mais divertidos, foi necessário fazer uma selecção daqueles a mapear numa aplicação computacional. Por um lado, os exercícios com a bola de ginástica eram os mais interessantes, visto já existirem no mercado alguns dispositivos com a funcionalidade de fazer um controlo virtual e que são fisicamente uma bola de dimensões semelhantes à referida. Um exemplo é a *Wii Bowling Ball*. No entanto, os pacientes já consideram que a realização destes exercícios se assemelham a jogos e, como tal, o nível de motivação já é elevado.

No extremo oposto, os que deixam os pacientes mais desmotivados são os exercícios na cama, pois são os que lhes provocam mais dor, visto existir uma maior superfície de contacto do corpo com o aparelho. Assim, a nível de motivação, onde eventualmente se verificaria uma maior distinção entre a fisioterapia tradicional e a que recorre à nossa plataforma, seria no mapeamento destes exercícios.

Os restantes exercícios apresentam poucas variantes e têm um mapeamento difícil, por exigirem por si só muito dos pacientes, como é o caso dos exercícios com a *fitball*, *squeezeball* e paralelas. Existem ainda alguns que apesar de serem fáceis de utilizar e mapear para a nova plataforma, têm pouca originalidade, como os exercícios na cadeira de rodas (e.g. *SmartWheel* [Asato93]).

Assim, o ideal será desenvolver uma forma mais divertida de realizar os exercícios na cama e com a bola de ginástica, conjugados entre si.

### **3.1 Cenários e sistemas**

Nesta secção iremos descrever o sistema a desenvolver e alguns cenários de interacção, com o objectivo de mostrar as potencialidades da nossa plataforma e permitir uma visualização de possíveis soluções futuras.

O sistema desenvolvido será uma plataforma com duas componentes: projecção superior na cama, e um ecrã frontal com um ambiente virtual controlado pela bola.

A primeira componente consistirá num projector de topo, que reflecte para a cama onde o paciente se encontra. A interacção será através do toque na cama, com recurso às mãos, braços ou tronco.

Já no que refere à componente da bola, esta será usada para controlar objectos no ecrã (i.e. ambiente virtual). Um exemplo poderá ser um jogo no qual o utilizador tem de alcançar ou desviar-se de objectos rodando a bola na direcção pretendida. Se quer deslocar-se para a esquerda, deverá rodar a bola para o lado esquerdo, por exemplo.



Figura 3 – Cenário *Croco Bed*.

Pretendemos com esta abordagem tornar o processo de reabilitação mais divertido, juntando elementos físicos da fisioterapia e elementos virtuais de jogos computacionais. Com esta sobreposição do real e virtual iremos melhorar a experiência dos utilizadores, motivando-os para a execução dos exercícios (mesmo os mais difíceis) e, consequentemente, aumentar o seu desempenho. Desta forma, estaremos não só a contribuir para uma experiência mais envolvente, como para um processo de fisioterapia mais eficiente e eficaz.

#### *Croco Bed*

O Francisco está a ter um péssimo dia. O tempo está mau e ele acordou nostálgico e triste. Mas as coisas parecem que finalmente vão melhorar! Quando chegou à fisioterapia, a terapeuta propôs-lhe que começasse pelos exercícios na *croco bed*. Embora o Francisco não goste dos exercícios na cama, porque lhe provocam muita dor, utilizar a *croco bed* é sempre um motivo de diversão (Figura 3). Na cama é projectada de forma aleatória imagens de crocodilos que se deslocam na sua direcção. Como o exercício seleccionado foi o enrolamento tradicional, pode alcançar o seu objectivo e matar muitos crocodilos simultaneamente. Para isso, basta-lhe bater nas suas cabeças, usando as mãos, ou rodando todo o corpo ou braços por cima deles. Quantos mais crocodilos forem mortos, e quanto maior for a velocidade da execução, maior será a pontuação acumulada. Será que vai melhorar a sua pontuação e subir no ranking?

#### *Rolling Balls*

Desta vez o Francisco não vê a hora de chegar ao centro de fisioterapia! A terapeuta prometeu deixá-lo experimentar a *rolling balls*. Embora já tenha visto os colegas a utilizarem-na, o Francisco ainda teve essa oportunidade. E eis que chegou o momento: no ecrã é projectado um cenário virtual em que o objecto principal é navegar em quatro direcções. O Francisco roda a bola de ginástica para um lado, rola para o outro, passando-a por cima das pernas, desvia-se dos obstáculos, e logo na sua primeira experiência com esta plataforma fica em 2.º lugar no ranking geral. Devido à velocidade de execução de movimentos e número de obstáculos dos quais se conseguiu

desviar, a sua pontuação foi quase óptima. Tem de voltar a tentar para melhorar o seu desempenho, alcançar o primeiro lugar e passar ao próximo nível de dificuldade.

#### 4. CONCLUSÕES

A fisioterapia, na sua forma tradicional, pode ser uma actividade aborrecida. Isto ganha especial relevância para aqueles que dependem dela diariamente. Para combater esta lacuna, propomos a utilização de um ambiente virtual durante o processo de reabilitação, que, ao contrário dos sistemas semelhantes já existentes, tenha um desenho informado. Este poderá desempenhar um papel importante na vida dos pacientes tetraplégicos, tornando a fisioterapia numa actividade mais aliciante. Porém, para alcançar este objectivo temos de perceber as necessidades reais destes pacientes.

Neste artigo é apresentando um estudo e caracterização detalhada do processo actual de fisioterapia, descrevendo os seus principais exercícios e limitações. Através do contacto constante com a população-alvo, apresentamos um conjunto de implicações para o desenho de uma plataforma tecnológica e propostas de desenho sob a forma de cenários de interacção.

#### 5. AGRADECIMENTOS

Os autores agradecem a disponibilidade das terapeutas Inês Ponte e Vanessa Jorge e dos pacientes do Lar CVP. Este trabalho foi financiado pela FCT através do programa PIDDAC. Tiago Guerreiro e Hugo Nicolau foram apoiados pela FCT, através das bolsas SFRH/BD/28110/2006 e SFRH/BD/46748/2008, respectivamente.

#### 6. REFERÊNCIAS

- [Andrade03] M. L. Andrade. Fisioterapia e informática, juntas na independência do portador de severo comprometimento neurológicos. *Technical Report*, 2003.
- [Asato93] K. Asato, R. Cooper, R. Robertson, J. Ster. Smartwheels: development and testing of a system for measuring manual wheelchair propulsion dynamics. *Technical Report*, 1993.
- [Faria06] F. Faria. Lesões vértebro-medulares – a perspectiva da reabilitação. *Revista Portuguesa de Pneumologia vol. XII no. 1*, 2006.
- [Foulds07] R. Foulds, D. Saxe, A. Joyce, S. Adamovich. Sensory-motor enhancement in a virtual therapeutic environment. 2007.
- [Kizony06] R. Kizony, P. Weiss, M. Shahar, D. Rand. Theragame: a home based virtual reality rehabilitation system. *Esbjerg, Denmark*, 2006.
- [Merino08] D. Merino, B. Chiarion, P. Pizzelli. Estudo do papel do fisioterapeuta nas principais complicações do traumatismo raquimedular na fase hospitalar: relato de caso clínico. 2008.
- [Rand04] R. Rand, D. Kizony, P. Weiss. Virtual reality rehabilitation for all: Vivid GX versus Sony Playstation II Eyetoy. 2004.
- [Sisto08] S. Sisto. Virtual reality rehabilitation and wii habilitation. *Technical Report*, 2008.



# Bases para Reabilitação Assistida

Rita Pereira    Tiago Guerreiro    Hugo Nicolau    Daniel Gonçalves    Joaquim Jorge

Technical University of Lisbon / INESC-ID

R. Alves Redol, 9, 1000-029

Lisbon, Portugal

rita.t.c.pereira@ist.utl.pt, {tjvg, hman}@vimmi.inesc-id.pt, {daniel.goncalves, jaj}@inesc-id.pt

## Sumário

O processo de fisioterapia consiste em devolver alguma qualidade de vida a deficientes motores, através do treino de um conjunto de movimentos. Cabe ao fisioterapeuta conseguir observar, interpretar e avaliar o estado actual e evolução dos seus pacientes, de forma a maximizar o seu desempenho físico. Neste artigo, apresentamos uma análise ao processo actual de fisioterapia, num centro de reabilitação para tetraplégicos, identificando as suas principais limitações e oportunidades para uma ferramenta tecnológica. Seguindo uma abordagem de desenho centrado no utilizador, é descrita uma plataforma de suporte aos fisioterapeutas, cujo principal objectivo é tornar a reabilitação num processo mais fiável e robusto. Avaliações preliminares com a população-alvo confirmam a utilidade da nossa abordagem, contribuindo para um acompanhamento mais preciso. Por fim, são apresentados alguns cenários de interacção ilustrando todas as potencialidades do sistema.

## Palavras-chave

*Movimentos, Reabilitação, Fisioterapia, Tetraplégicos, Reprodução do Movimento, Captura do Movimento*

## 1. INTRODUÇÃO

Nas últimas décadas, o avanço da tecnologia tem vindo a melhorar a qualidade de vida das pessoas portadoras de deficiência. No entanto, os desenvolvimentos práticos ainda são escassos para os tetraplégicos e, em particular, para o seu processo de reabilitação.

A fisioterapia é uma ciência na área da saúde voltada para o entendimento da funcionalidade humana que actua nas mais diferentes áreas com procedimentos, técnicas, metodologias e abordagens com o objectivo de avaliar, tratar, minimizar e prevenir as mais variadas disfunções. Este processo, quando aplicado a tetraplégicos, visa sobretudo que este ganhe independência nos movimentos que realiza e que se adapte positivamente à sua condição de vida [Bromley06].



**Figura 1 – Reabilitação tradicional (Análise de Tarefas)**

Ao fisioterapeuta cabe o papel de observar, interpretar e dirigir a acção para otimizar a resposta do paciente. No entanto, estes processos terapêuticos são longos e implicam grande dedicação e motivação por parte do mesmo.

Consistem ainda na realização de padrões de movimento repetitivos durante as sessões, o que pode ser um factor de desmotivação para os pacientes [Bromley06].

Este trabalho tem como objectivo desenvolver ferramentas de auxílio a fisioterapeutas e, consequentemente, melhorar o processo de reabilitação. Começámos por estudar as rotinas diárias, sessões e procedimentos num centro de reabilitação, apresentando, neste artigo, um conjunto de limitações e exigências de uma solução tecnológica. O nosso objectivo é proporcionar aos fisioterapeutas um sistema computacional, com mecanismos eficientes e precisos, para o acompanhamento dos seus pacientes. Estes mecanismos incluem: guardar todos os exercícios que o paciente realiza para posterior avaliação e visualização; ter uma medida precisa do paciente e da sua evolução; e poder comparar o desempenho de duas sessões distintas no tempo.

Neste artigo, apresentamos os principais resultados da análise das tarefas realizada com os principais intervenientes no processo de reabilitação, assim como um conjunto de requisitos para uma ferramenta de suporte. Apresentamos ainda um trabalho em desenvolvimento, de uma plataforma virtual de reabilitação.

## 2. PERCEBER A FISIOTERAPIA

Para entendermos melhor o processo de reabilitação, realizamos um estudo de 10 dias num centro de reabilitação com os principais utilizadores deste sistema, ou seja, os fisioterapeutas, e pacientes (Figura 1). Apesar de acreditarmos que o nosso trabalho poderá ser valioso para a população em geral, concentramos os nossos estudos em pacientes tetraplégicos. A escolha desta população deve-

se essencialmente à longa duração do tratamento que poderá se durar entre vários meses a alguns anos. Assim, para que o processo de reabilitação e acompanhamento seja minimamente eficaz é necessária uma boa análise, armazenamento e noção da evolução dos pacientes.

Com vista a alcançar este objectivo, foi realizado um estudo preliminar com três fisioterapeutas e sete pacientes tetraplégicos. A análise consistiu em entrevistas e questionários com todas as partes intervenientes, e uma observação minuciosa dos exercícios da fisioterapia. As entrevistas foram realizadas após as sessões de fisioterapia e ajudou-nos a compreender os objectivos, limitações e dificuldades de cada método. Deste período de observação destacamos as seguintes conclusões:

- f A maioria dos exercícios é normalmente realizada junto do fisioterapeuta. Durante a sua execução, o fisioterapeuta tira notas e observa como é que o movimento está a ser executado.
- f Alguns exercícios são realizados na própria cadeira de rodas (por exemplo, mover um braço), mas, para ser executado correctamente, depende de um conjunto de restrições (por exemplo, manter o tronco firme).
- f A execução repetitiva de exercícios errados, não respeitando as restrições anteriormente identificadas, pode ter um efeito perigoso ou até mesmo regressivo sobre a reabilitação do paciente.
- f É difícil o fisioterapeuta ter uma visão completa do movimento quando este precisa de intervir na sua realização (Figura 1).
- f A análise dos movimentos observados pelo fisioterapeuta é registada com um valor aproximado. Este valor pode variar de um fisioterapeuta para o outro. Além disso, considerando a duração do processo, é impossível garantir a coerência entre as avaliações prejudicando assim o registo da evolução do paciente;
- f Os pacientes não têm qualquer retorno visual sobre o seu movimento ou desempenho (e.g. distância para um movimento preciso).

A partir da análise, consideramos que uma plataforma informática suportada com um sistema de monitorização preciso é uma adição valiosa para os procedimentos de reabilitação. No sistema proposto, é importante destacar tantos os pacientes como os fisioterapeutas como população-alvo. Para os fisioterapeutas, este sistema trará benefícios como a partilha de informação, padrões de movimento, análise e visualização do movimento; para os pacientes, poderá aumentar a motivação para alcançar os seus objectivos e, assim, melhorar o seu desempenho nos exercícios propostos.

### 3. REABILITAÇÃO VIRTUAL ASSISTIDA POR COMPUTADOR

Após o nosso estudo preliminar foi-nos possível identificar as principais falhas e limitações no processo de reabilitação, concluindo que uma plataforma de suporte poderia ser benéfica para ambas as partes envolvidas (i.e. pacientes e fisioterapeutas). Assim, foi criado um sistema de fisioterapia assistida que tem como base os seguintes requisitos:

- f *Persistência de Dados:* Todos os dados devem ser persistentes e coerentes, para que possam ser visualizados e partilhados entre fisioterapeutas.
- f *Captura de Movimentos:* Deve ser possível gravar o movimento realizado para posterior análise e reprodução;
- f *Precisão:* A plataforma deverá permitir obter a noção exacta e precisa de aspectos relacionados com o movimento, por exemplo, o alcance da mão do paciente;
- f *Reprodução do Movimento:* Deve ser possível reproduzir o movimento a qualquer instante, para análise e avaliação do mesmo;
- f *Comparação do Movimento:* Deve ser possível reproduzir dois movimentos sobrepostos, para que possam ser comparados (e.g. avaliação da evolução);
- f *Extracção Automática de Informação:* Deve ser capaz de enriquecer a visualização e facilitar a análise dos exercícios (e.g. apresentar automaticamente a distancia entre dois pontos na comparação de um movimento particular).
- f *Fácil Configuração:* O terapeuta deverá ser capaz de preparar um exercício com pouco esforço e sem conhecimentos técnicos específicos da tecnologia usada.

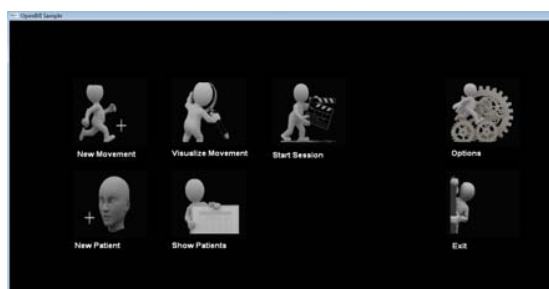
Para cumprir os objectivos e garantir que as exigências sejam cumpridas, a nossa abordagem usa um sistema de seguimento virtual (i.e. *tracking*) com marcadores ópticos. Este seguimento é realizado através da reflexão proporcionada pelos marcadores colocados no corpo do utilizador [Zhou04] (Figura 2).

A escolha deste método de captura de movimentos deve-se principalmente à sua precisão.

No entanto, a posição de cada um dos pontos é independente dos restantes, não existindo qualquer relação entre eles. Logo, para conseguirmos manter um constante seguimento de cada marcador, mesmo quando existe um cruzamento, optámos por usar Filtros de *Kalman* [Welch 95]. Estes são caracterizados por usarem um modelo de um sistema dinâmico (i.e. com base nas leis físicas do movimento), onde são conhecidas as entradas de controlo e as medições para formar uma estimativa das quantidades variáveis do sistema (i.e. o seu estado).



Figura 2 – Exemplo do sistema de seguimento com base em marcadores



**Figura 3 – Menu Principal da aplicação**

O seguimento é efectuado a vários pontos distintos, em simultâneo, alguns deles directamente relacionados (e.g. dois pontos no braço, como na Figura 2) e outros com relações indirectas (e.g. tronco e braços). É importante notar que apesar de se manter uma noção interna do esqueleto e onde é colocado cada marcador, não são usados corpos rígidos. Ou seja, os marcadores são pontos isolados no espaço, proporcionando aos fisioterapeutas uma maior flexibilidade na fase de montagem do sistema.

O filtro de Kalman tem duas fases distintas: previsão e actualização. A fase de previsão utiliza a estimativa do estado no instante anterior para produzir uma estimativa do estado na fase actual. Na fase de actualização, a predição "*a priori*" é combinada com informações actuais da observação, para refinar a estimativa do estado.

No entanto, este valor não é suficiente para identificarmos todos os pontos recebidos. Para atingir este objectivo, precisamos de usar a distância de *Mahalanobis* [Pinho05] e determinar qual dos pontos medidos é o mais próximo para a predição calculada.

Este processo é repetido a cada passo e a sua nova estimativa é usada na iteração seguinte.

#### 4. PLATAFORMA DE REABILITAÇÃO

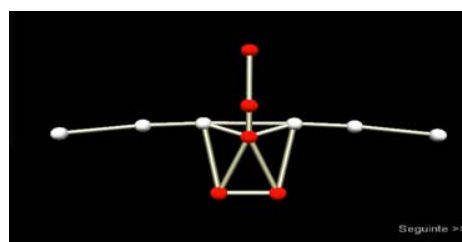
A plataforma de suporte permite aos terapeutas gerir todas as informações dos seus pacientes e respectivas sessões de fisioterapia, mantendo um histórico dos exercícios realizados (Figura 3). Além disso, permite-lhes comparar os dados entre sessões, ou mesmo entre pacientes. As funcionalidades mais relevantes irão ser descritas em seguida.

##### 4.1 Gravar movimento

A plataforma permite que o fisioterapeuta consiga gravar um movimento para posterior visualização ou comparação e, em ambos os casos, escolher quais os pontos mais



**Figura 4 – Utilizador com os marcadores**



**Figura 5 – Selecção de pontos no esqueleto**

relevantes para a avaliação do movimento (Figura 4).

De seguida poderá selecciona-los no esqueleto tridimensional (Figura 5) e, assim, começar a gravação.

#### 4.2 Reprodução do movimento

Após a gravação, o fisioterapeuta pode optar por reproduzir, navegar ou visualizar com detalhe o movimento, numa vista tridimensional (Figura 6). A aplicação permite ainda analisar os pontos escolhidos e observar as distâncias e ângulos entre os membros. Esta função comporta-se como um *player*, onde podemos fazer Iniciar, Pausar ou até mesmo alterar a velocidade do movimento.

#### 4.3 Comparação de movimentos

A qualquer momento, o fisioterapeuta pode seleccionar mais do que um movimento para compará-los. Para facilitar esta tarefa, são apresentadas duas barras temporais distintas (Figura 7), permitindo um controlo manual sobre as reproduções dos movimentos. Estamos actualmente a trabalhar numa sincronização automática para auxiliar os terapeutas a encontrar mais facilmente os mesmos instantes no movimento.

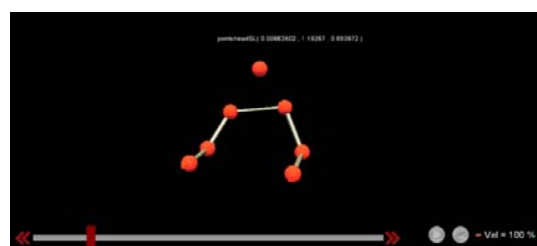
#### 5. AVALIAÇÃO PRELIMINAR

Foi realizada uma avaliação preliminar com dois fisioterapeutas, onde podemos concluir que, com a utilização do sistema, as tarefas podem ser realizadas de um modo rápido, e que toda a informação é obtida no instante da execução do movimento. No fim desta avaliação, os fisioterapeutas reconheceram que a com o sistema, conseguem obter a informação do movimento de um modo rápido e com maior precisão. No geral gostaram do sistema e acreditam que seja uma melhoria significativa para auxiliar as suas actividades diárias.

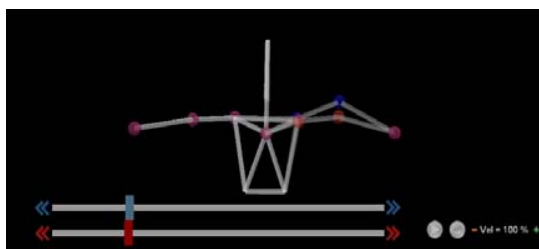
#### 6. CENÁRIOS

As capacidades da plataforma podem ser ilustradas com os seguintes cenários:

##### 6.1 Captura do movimento



**Figura 6 – Reprodução do Movimento**



**Figure 7 – Comparação entre dois movimentos**

O José, como em todas as quartas feiras, tem fisioterapia às 9:30 com o Pedro. Uns minutos depois de chegar, o fisioterapeuta Pedro, começou por colocar os marcadores de monitorização no seu paciente. Após terminar esta tarefa, o Pedro iniciou a aplicação e pediu ao José que colocasse os braços apoiados na cadeira. Começou então por escolher na aplicação a opção “Gravar”, onde de seguida lhe foi apresentada uma lista de pacientes que frequentam o centro. Escolheu o paciente José Martins, carregou em seguinte e apareceu-lhe uma lista de todos os movimentos executados pelo José. O Pedro decidiu iniciar uma nova sessão e escolheu a opção “Começar a Gravar”. Seleccionou o movimento “Levantar o Braço Esquerdo” e foi-lhe apresentado um esqueleto, para que fossem seleccionados os marcadores que foram colocados no José. Após todos os pontos introduzidos, o Pedro deu sinal ao paciente que iria começar a gravar, carregando no respectivo botão. Enquanto o José realizava o exercício, o esqueleto tridimensional na aplicação movia-se de acordo com os seus movimentos. Quando terminou o exercício o Pedro carregou em “Terminar” e “Guardar Movimento”.

## 6.2 Reprodução do movimento

A Maria está pronta para começar mais uma sessão de fisioterapia. Com o auxílio da Vanessa consegue-se equilibrar, sentada, em cima da cama. No entanto, na última sessão a Vanessa tinha notado que algo estava errado com o movimento da sua paciente. Hoje, decide visualizar o exercício em detalhe e mostra à Maria o que estava a fazer de mal. Pressiona o botão “Visualizar” e é mostrada uma barra, com as opções de “Iniciar”, “Ajustar Velocidade” e “Parar”. A Vanessa decide reduzir a velocidade em 50% e carrega em “Iniciar”. No momento em que a Maria está a começar a inclinar-se, erradamente, ela pára a reprodução, explicando à Maria onde estava o erro. Desta forma, a Maria conseguiu facilmente perceber o que estava a fazer mal e corrigir o seu movimento.

## 6.3 Comparar movimentos

Depois de um longo dia de trabalho, o Pedro pretende realizar um balanço de como têm corrido as sessões de fisioterapia, e quais as alterações que devem ser aplicadas no plano de trabalho de cada paciente. Para isso, começou por analisar as suas evoluções desde a primeira sessão. Iniciou a aplicação, e escolheu a opção “Lista de Sessões”, carregou em seguinte e apareceu a lista de todos os pacientes. Seleccionou o André Lopes e de seguida apareceu uma listagem de todos os movimentos

gravados. O Pedro estava particularmente interessado no exercício de “Elevação do Braço Esquerdo”. Após seleccioná-lo, escolheu a opção “Comparar Movimentos”. Ao aparecer a listagem de todas as gravações do André com este movimento, o Pedro seleccionou os que tinham sido executados nos dias 12/12/2009 e 12/04/2010, carregando no botão “Comparar”. Foram apresentadas duas barras, uma para cada execução, um esqueleto tridimensional, e um botão de “Iniciar” e “Parar”. Ao carregar no “Iniciar”, notou que o André mostrou melhorias na extensão do braço, atingindo já um ângulo de quase 90° com o tronco.

## 7. CONCLUSÕES E TRABALHO FUTURO

Neste artigo apresentamos os nossos estudos preliminares, ao processo de reabilitação actual, em particular, o modo como os fisioterapeutas observam e avaliam o estado físico dos seus pacientes.

Propomos uma plataforma virtual de suporte e monitorização de movimento que permite aos terapeutas terem um acesso fácil e rápido a toda a informação sobre os exercícios realizados pelos pacientes: desempenho, evolução, e histórico. Como trabalho futuro iremos avaliar o nosso sistema, junto de todos os intervenientes da reabilitação (fisioterapeutas e deficientes motores), de forma a validar a nossa abordagem. Apesar de o nosso foco estar actualmente no suporte ao fisioterapeuta e em garantir um processo mais fiável e robusto, pretendemos no futuro melhorar a experiência do paciente tornando-a mais divertida (e.g. através de jogos).

## 8. AGRADECIMENTOS

O trabalho foi financiado pela FCT através do programa PIDDAC. Tiago Guerreiro e Hugo Nicolau foram apoiados pela FCT, através das bolsas SFRH/BD/28110 /2006 e SFRH/BD/46748/2008, respectivamente.

## 9. REFERÊNCIAS

- [Zhou04] Zhou, H. & Hu, H, A Survey – Human Movement Tracking and Stroke Rehabilitation, Technical Report, 2004.
- [Arulampalam02] Arulampalam M., Maskell S., Gordon N., Clapp T. ,A Tutorial on Particle Filters for On-line Nonlinear/Non-Gaussian Bayesian Tracking, IEEE Transactions on Signal Processing, vol. 50, pp. 174/188, 2002
- [Welch95] Welch G., Bishop G., An Introduction to Kalman Filter, University of North Carolina at Chapel Hill, 1995
- [Pinho05] Pinho, R. R.; Tavares, J. M. R. S. & Correia, M. VA Movement Tracking Management Model with Kalman Filtering, Global Optimization Techniques and Mahalanobis Distance, ICCMSE 2005 - International Conference on Computational Methods in Science and Engineering, 2005
- [Bromley06] Tetraplegia and paraplegia : a guide for physiotherapists. - 6th ed. - I. Bromley. Edinburgh : Churchill Livingstone, 2006

# Psicoterapia Georreferenciada com Dispositivos Móveis

Marco de Sá, Luís Carriço,  
Ricardo Pereira  
LaSIGE & Univ. de Lisboa  
{marcosa,lmc, repeira}@di.fc  
.ul.pt

João Faria, Isabel Sá  
Univ. de Lisboa  
Lisboa  
jnunofaria@gmail.com,  
misa@fp.ul.pt

Gustavo Zurita, Nelson Baloian,  
Universidad del Chile  
gzuri-  
ta@fen.uchile.cl, nbaloian@gmail.co  
m

## Sumário

*O treino de competências sociais, como parte da psicoterapia, para crianças e adolescentes, requer que estes participem em actividades exteriores, nas quais têm que completar tarefas, tais como falar com alguém ou visitar um local específico. De momento, a incapacidade, por parte dos terapeutas, de monitorizar os seus pacientes, promover esforços colaborativos e fortalecer atitudes positivas, é uma questão importante que afecta tanto o processo da terapia como os seus resultados.*

*Neste artigo detalhamos a avaliação de um protótipo para um sistema colaborativo móvel e georreferenciado que suporta terapia de grupo in-situ. O sistema visa fornecer meios ao terapeutas, para monitorizar os seus pacientes, as suas localizações e os seus feitos, e inclui mecanismos de comunicação que facilitam a cooperação entre pacientes e terapeuta. Descrevemos o conceito por detrás do projecto, os protótipos de baixa fidelidade e os testes que foram levados a cabo para validá-los. São discutidos resultados e o é esboçado o trabalho futuro.*

## Palavras-chave

*Colaboração Georreferenciada, Dispositivos Móveis, Psicoterapia.*

## 1. INTRODUÇÃO

Por variadas razões, as pessoas sentem-se, frequentemente, desconfortáveis quando confrontadas com situações sociais comuns, ao ponto de, em certos casos, prejudicarem a sua vida diária [11]. Este tipo de problema afecta todas as idades, sexos e culturas sendo, contudo, mais visível em crianças. Casos mais graves, como Autismo e Síndrome de Asperger, são também exemplos comuns de perturbações com sintomas semelhantes [5]. Consequentemente, quando confrontadas com este tipo de problemas, as crianças e adolescentes são geralmente acompanhadas por terapeutas, participando em treinos de competências sociais.

O treino de competências sociais (TCS) é um tipo de procedimento de psicoterapia no qual os pacientes geralmente têm de executar um conjunto de tarefas individualmente (por exemplo, pedir um café num bar; dar um elogio a alguém) ou em grupo (por exemplo, falar uns com os outros; trabalhar em equipa) [5], trabalhando assim com os seus colegas e pais [4]. De momento, este processo é geralmente síncrono, requerendo que o terapeuta determine, durante uma sessão, um conjunto de tarefas para os pacientes executarem em casa (com os seus pais) ou habitualmente na escola (com os seus colegas). As tarefas de campo, embora realizadas individualmente, são preferencialmente empreendidas como um grupo, para que a seja mantida a motivação mesmo quando o terapeuta não está acessível. As tarefas são, habitualmente, definidas como um conjunto de

actividades que podem ser realizadas através de uma rota pré-definida que os pacientes seguem, visitando locais específicos onde lhes são propostos desafios. Cada vez que um paciente consegue realizar uma tarefa, anota os seus/seus pensamentos e regista um conjunto de valores para sentimentos como timidez, stress, ansiedade e dificuldade.

No final, os resultados são partilhados com o terapeuta durante uma sessão, na qual a experiência é discutida e os sentimentos são explicados. Questionários de papel e anotações são usadas para recolher dados enquanto no local, e suportar a troca de informação com o terapeuta.

Tendo em conta o meio utilizado, existem vários inconvenientes. Em primeiro lugar, como o controlo por parte do terapeuta é intermitente, os pacientes desistem frequentemente do processo da terapia e fornecem, muitas vezes, dados e respostas falsas. Adicionalmente, a cooperação entre os vários actores é dificultada pela usual falta de comunicação e capacidades sociais; a motivação é extremamente baixa enquanto longe do consultório do terapeuta, do controlo dos pais e, especialmente, quando a cooperação entre pacientes não é conseguida.

Neste artigo apresentamos um trabalho que visa avaliar a viabilidade, adequação e aceitação de uma ferramenta colaborativa móvel e georreferenciada para suportar o TCS. É discutido o processo de desenho e são apresentados os protótipos iniciais, salientando também o processo de validação que foi conduzido e os resultados iniciais. São também apresentadas direcções futuras.

## 2. TRABALHO RELACIONADO

Avanços recentes na tecnologia mostraram que o uso da mesma, aplicada à psicoterapia, pode melhorar o trabalho dos terapeutas [7]. Em particular, o uso de dispositivos móveis tem ganho ímpeto e várias ferramentas têm sido desenvolvidas [8], a maioria utilizando dispositivos portáteis como PDAs ou TabletPCs para terapias de auto-controlo ou procedimentos de relaxamento.

Ainda assim, a cooperação e o apoio a actividades de grupo são ainda muito escassos. Noutros domínios, alguns sistemas destinados a questões semelhantes têm sido desenvolvidos, abordando, no entanto, aspectos relacionados com a necessidade de apoio à comunicação entre vários dispositivos, portáteis e até mesmo não móveis [9][10]. Contudo, embora o objectivo habitual seja suportar o trabalho cooperativo num ambiente espacial distribuído, gerindo a informação que é partilhada entre dispositivos, a especificidade do cenário da terapia de grupo requer um cuidado especial. Em domínios como este, dar ênfase significativo à facilidade de uso, à adequação ao procedimentos clínicos e à especificidade dos artefactos usados [6], é da maior importância. Para além disso, e principalmente, para o TCS, é essencial permitir que os terapeutas mantenham a percepção da localização e actividades dos seus pacientes, não só para controlar o processo mas também para que sejam capazes de intervir sempre que necessário.

## 3. BACKGROUND E PROCESSO DE DESENHO

Este projecto surgiu como a conjugação de dois sistemas diferentes com propósitos distintos. O primeiro visa apoiar o trabalho geo-colaborativo e criação de conhecimento, e apoiar anotações georreferenciadas (em diversos formatos) que podem ser monitorizadas e partilhadas em tempo real através de dispositivos móveis [3]. Os objectivos do segundo projecto são o apoio a práticas de terapia comportamental cognitiva fornecendo meios para que terapeutas e pacientes completem as suas actividades (por exemplo, registo de pensamentos, classificação de sentimentos) de forma ubíqua, também utilizando dispositivos móveis [2].

No decorrer de alguns testes com a ferramenta de psicoterapia em grupo, na qual vários terapeutas estiveram envolvidos, a sua necessidade de poder conduzir sessões fora do consultório e monitorizar os seus pacientes enquanto estes completam as suas actividades ao ar livre, foi notória e muitas vezes mencionada. Portanto, a ideia de combinar estes dois sistemas, que permitem aos terapeutas partilhar um conjunto de artefactos de colecta de dados com os seus pacientes, num sistema que pode fornecer a percepção do local de vários utilizadores e comunicar com os mesmos em tempo real, foi rapidamente aceite de forma entusiástica por todos os terapeutas envolvidos.

### 3.1 Objectivos do Sistema

Baseado nas necessidades que os terapeutas de TCS apresentaram, o sistema visa atender a várias questões e melhorar o processo: (1) fornecendo uma plataforma colaborativa integrada tanto para terapeutas como para

pacientes; (2) apoiando a monitorização dos pacientes enquanto realizam as tarefas, em tempo real; (3) oferecendo mecanismos de comunicação entre todos os participantes; (4) incluindo registos multimodais de experiências e (5) promovendo motivação colaborativa entre pacientes e terapeuta. Isto pode ser conseguido através da fusão de um conjunto de características colaborativas georreferenciadas, herdadas do primeiro projecto com artefactos de terapia melhorados, que suportam mecanismos de revisão e a colecção de informação in-situ, oferecidos pelo segundo.

No geral, isto resulta num sistema que, por um lado, permite aos utilizadores (pacientes) aceder e criar informação (por exemplo, vídeo/áudio/anotações de texto) que é georreferenciada, utilizando um mapa digital, directamente nos seus dispositivos móveis (por exemplo, smartphone, PDA); e por outro lado, permite ao terapeuta gerir toda a informação que é gerada, aceder à localização de todos os utilizadores, gerir e distribuir informação pelos os participantes. Deste modo, o sistema visionado substitui os tradicionais questionários de papel por um dispositivo móvel, que acompanha o paciente ao longo do processo de TCS, integrando o terapeuta no processo, mesmo fora do consultório.

### 3.2 Funcionalidades

#### 3.2.1 Criação de dados georreferenciados

De uma perspectiva geral, o sistema de geração de dados georreferenciados fornece mecanismos para os utilizadores criarem informação georreferenciada, seleccionando uma localização no mapa e criando anotações estruturadas, textuais ou baseadas em esboços [3]. O sistema também inclui um mecanismo de comunicação subjacente, baseado em serviços web. Com este mecanismo, a informação gerada pode ser distribuída por vários utilizadores, que podem procurar pela informação através da sua localização, directamente num mapa.

#### 3.2.2 Recolha Multimodal de Dados de Terapia

A plataforma de psicoterapia inclui um conjunto de funcionalidades que permitem aos terapeutas criar artefactos específicos para um paciente, permitindo a recolha de dados usando varias modalidades (por exemplo, voz, vídeo, imagem, texto) [2]. Cada artefacto pode ser utilizado num dispositivo móvel. Os dados, juntamente com logs de utilização, são armazenados localmente em ficheiros XML e podem ser sincronizados com outros dispositivos.

### 3.3 Processo de Desenho

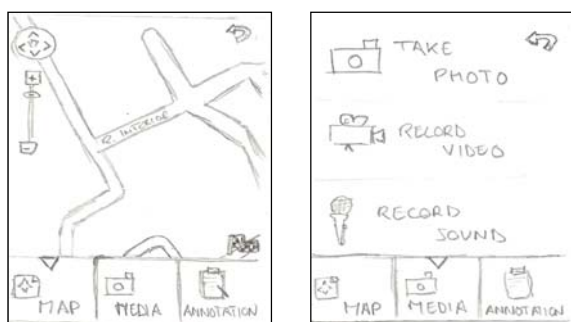
O desenho do sistema seguiu uma abordagem centrada no utilizador [1], na qual as várias fases foram acompanhadas de perto por utilizadores finais e o seu contributo foi fornecido desde o início. Como ponto de partida, foram realizadas reuniões com terapeutas e membros dos dois projectos, resultando no sistema em desenvolvimento. A definição de requisitos foi baseada nas funcionalidades existentes, pedidos dos terapeutas e descrições das necessidades dos pacientes. Foram definidos alguns cenários e feitas algumas simulações para identificar questões e requisitos adicionais.



Estes requisitos foram cuidadosamente discutidos com terapeutas e investigadores de psicoterapia e, uma vez chegado a um consenso, o conceito do projecto evoluiu para a fase de prototipagem.

### 3.4 Protótipos

Para testar o conceito, foram criados dois protótipos de baixa fidelidade. Um para a aplicação do terapeuta, desenhado para um TabletPC/Slate e outro para a ferramenta dos pacientes, direccionada a smartphones. O processo de criação destes protótipos seguiu uma metodologia avançada, especificamente concebida para apoiar o desenho de interacção móvel [1]. A figura 1. mostra dois esboços para a aplicação do paciente. No ecrã da esquerda os pacientes podem acompanhar as suas tarefas, localização e escolher rapidamente uma opção de meio de captação para criar uma anotação. O ecrã da direita exibe os diferentes tipos de media disponíveis. Os pacientes podem tirar uma foto, gravar um vídeo, ou anotação áudio, e enviá-lo imediatamente ao terapeuta e a outros pacientes (se permitido pelo terapeuta).



**Figura 1. Protótipos de baixa fidelidade para a ferramenta do paciente.**

As figuras 2 e 3 mostram a aplicação do terapeuta. Aqui, o terapeuta é capaz de monitorizar a localização dos pacientes bem como as suas movimentações e definir novas tarefas que são automaticamente actualizadas na rota dos pacientes. O menu mais abaixo permite ao terapeuta gerir as variadas opções e mandar mensagens globais, directamente a todos os pacientes. Na tela da direita está sempre disponível a informação individual de cada participante, incluindo tarefas completadas, sentimentos e sua intensidade, e o progresso actual.



**Figura 2. Esboço da ferramenta do terapeuta para TabletPC.**

Adicionalmente, estão disponíveis atalhos para a informação de cada paciente (por exemplo, anotações, fotos, perfil), juntamente com opções de mensagens individuais. A figura 2 mostra a caixa de texto na qual as mensagens podem ser lidas, escritas e adicionados ou removidos destinatários seleccionando as suas fotos. A figura 3 apresenta o ecrã da interface de utilizador do terapeuta, mostrando um mapa no qual se podem gerir as varias localizações dos desafios pré-estabelecidos, bem como as posições e movimentos dos pacientes.

### 4. AVALIAÇÃO INICIAL

Para validar os dois protótipos, foi conduzida uma experiência inicial do tipo Wizard-of-Oz. Os protótipos de baixa fidelidade foram usados por um terapeuta e por dois pacientes simulados, no campos de uma universidade. Os pacientes simulados eram jovens estudantes, sem historial de terapia, um rapaz e uma rapariga, escolhidos aleatoriamente. O terapeuta tem vasta experiência em TCS e não esteve envolvido em nenhum dos projectos anteriores. Foram emprestados lápis aos participantes para anotar (sempre que quisessem criar anotações textuais) e marcar qualquer informação necessária directamente nos esboços.

Foi atribuído um mediador ("Wizard") por cada participante, acompanhando o utilizador e agindo quer como a aplicação, quer como mecanismo de comunicação. Para apoiar a comunicação entre os vários mediadores, foram utilizados telemóveis e foram efectuadas chamadas para partilhar a localização de cada participante. O mediador do terapeuta marcava então, no mapa da aplicação, a localização de cada paciente simulado. As anotações foram trocadas por SMS ou MMS se incluíssem imagens, vídeo ou som.



**Figura 3. Protótipo do terapeuta durante a experiência de campo.**

Para que o teste fosse realizado, o terapeuta definiu três actividades, ao longo de seis locais diferentes, explicadas a cada participante durante uma breve introdução:

- Pedir a alguém, no bar, para passar à frente na fila.
- Perguntar no bar o preço de um café e comprá-lo.
- Pedir permissão para se sentar numa mesa ocupada.

A experiência demorou cerca de uma hora. No final, foram efectuadas entrevistas curtas individualmente e posteriormente foi realizada uma reunião de grupo para comentários e sugestões.

### 5. RESULTADOS E DISCUSSÃO

Os resultados da experiência foram bastante positivos. No que toca à viabilidade e eficácia de todo sistema, as

características visionadas foram muito apreciadas pelo terapeuta. A capacidade de aceder à localização dos pacientes, incluindo o tempo gasto em tarefas e pontos específicos, foi considerada bastante útil. Para além de permitir ao terapeuta controlar a rota dos pacientes, por razões de segurança (dado que a maioria dos pacientes são crianças), permite também controlar se o tempo gasto num local foi suficiente para concluir uma tarefa, ou até mesmo para controlar se o paciente está a seguir a sua rota. Segundo o terapeuta participante, isto deu a oportunidade de recolher informação real, eficaz e fidedigna e deu também muito menos espaço para erro. Adicionalmente, isto também confere ao terapeuta a capacidade de enviar mensagens encorajadoras em momentos cruciais, quando os pacientes se despendem do processo ou se sentem mais stressados.

As opções de comunicação também foram apreciadas pelos terapeutas, especialmente a possibilidade de partilhar ou dissimular feitos de um paciente para outros. Todavia, surgiram algumas sugestões e melhoramentos. Em particular, a inclusão de um botão de S.O.S na aplicação do paciente e a possibilidade de permitir que pacientes trocassem mensagens entre eles. Os participantes que actuaram como pacientes também forneceram um contributo valioso. Ambos os participantes acharam que a ferramenta era bastante fácil de usar e permitiu que rapidamente compreendessem as suas tarefas. Seguir a rota pré-definida no dispositivo móvel, e reunir informação em diferentes formatos também foi uma tarefa agradável e simples de executar.

Novamente, foram feitas sugestões. Um participante sugeriu a inclusão de uma lista de sentimentos que pudessem ser seleccionados e quantificados, ao invés de escrevê-los. O segundo participante mencionou que, se as tarefas fossem subdivididas, seria possível ver mais tarefas completadas, aumentando então a motivação, e ajudando na definição de objectivos e prioridades. O terapeuta concordou com esta sugestão. Durante a experiência, foi também notório que anotações de voz e fotografias foram frequentemente usados e preferidos em relação às anotações textuais/descrições de feitos, já que permitem fornecer e receber, rapidamente, *feedback* do terapeuta, aumentando a confiança ao longo do processo.

O uso de dispositivos móveis foi também considerado muito adequado visto que permite, aos pacientes, passar despercebidos enquanto recolhem informação e a partilham com o terapeuta. O terapeuta também concordou e considerou todo a infra-estrutura muito apropriada para estas situações, visto que a maioria dos pacientes já possuem os seus próprios dispositivos.

Em suma, o sistema foi considerado muito útil e fácil de utilizar. Da perspectiva do terapeuta, seria altamente benéfico aplicá-lo em TCS com crianças e adolescentes. A adição da comunicação entre pacientes, permitiria que se aumentasse a cooperação e, conseqüentemente, a motivação, resultados e adesão à terapia.

## 6. CONCLUSÕES E TRABALHO FUTURO

Neste artigo apresentámos o estudo de validação para o conceito de um sistema de terapia de grupo colaborativa

in-situ e georreferenciado. O conceito deste projecto é baseado em dois sistemas existentes e funde funcionalidades dos mesmos, de forma a permitir que terapeutas possam gerir pacientes que circulam ao longo de localizações diferentes mas que colaboram de forma a atingir objectivos específicos e ao mesmo tempo motivando-se entre si. Para testar a sua viabilidade, foi criado um conjunto de protótipos de baixa fidelidade e efectuada uma experiência do tipo Wizard-of-Oz. Um terapeuta e pacientes simulados participaram em sessões testando as duas vertentes e extensões que compõem o sistema. Globalmente, os resultados foram significativamente positivos e o terapeuta participante apreciou o conceito e encorajou vivamente a sua continuação para uma futura fase de desenvolvimento. Além disso, os restantes participantes, actuando como pacientes, também ficaram muito satisfeitos com a experiência, afirmando até que usariam o sistema para diversos fins do seu interesse. Com base nestes resultados, serão conduzidas novas experiências, com grupos maiores e terapeutas adicionais, usando protótipos novos e melhorados, e serão testadas novas funcionalidades. Está também a ser desenvolvido um protótipo de software que será testado em cenários reais, com crianças submetidas a este tipo de terapia.

## 7. REFERÊNCIAS

- [1] de Sá, M., Carriço, L. (2008). Lessons from Early Stages Design of Mobile Applications. *Procs of MobileHCI'08*, Amsterdam, September, pp. 127-136..
- [2] de Sá, M., et al. (2007). Ubiquitous Psychotherapy. *IEEE Pervasive Computing*, vol. 6 (1), pp. 20-27.
- [3] Zurita, G., Baloian, N., Peña, G., (2009) A Mobile Collaborative Visual System to Support Knowledge Creation. . *ICKM 2009*, China, December.
- [4] Bass, J. & Mulick, J. (2007). Social Play skill enhancement of children with autism using peers and siblings as therapists. *Psychology in the Schools*, 44(7), pp. 727-735.
- [5] White, S.; Keonig, K. & Scahill, L. (2007). Social Skills Development in Children. *Journal of Autism and Developmental Disorders*. 37.
- [6] Carriço, L., et al (2007). Managing Group Therapy through Multiple Devices. In *Procs of HCI International 2007*, July, vol. 4553/2007, pp. 427-436.
- [7] Garrard, C., (2000). Can computers improve the way doctors work? *Schweitz Med Wochenschr*, 130 (42).
- [8] Grasso, M. A., (2004). Clinical Applications of Hand Held Computing. 17th IEEE Symposium on Computer Based Medical Systems, 141-146, 2004.
- [9] Myers, B.A. (2000) The Pebbles Project: Using PCs and Hand-held Computers Together. *CHI'00*.
- [10] Geyer, W. et al.(2001) A Team Collaboration Space Supporting Capture and Access of Virtual Meetings. In *Procs. of GROUP'01*, Sept. 30-Oct. 3, USA.
- [11] Marchant, M., et al. (2007). Modifying socially withdrawn behavior. *Psychology in the Schools*, 44(8), pp. 779-794.

# Terapia do Medo para Crianças: Desenho e Prototipagem de uma Ferramenta para Dispositivos Móveis

Marco de Sá, Luís Carriço, Pedro Bernardo  
LaSIGE & Univ. de Lisboa  
{marcosa, lmc}@di.fc.ul.pt,  
pbernardo@lasige.di.fc.ul.pt

João Faria, Isabel Sá  
Faculdade de Psicologia, Univ. de Lisboa  
jnunofaria@gmail.com, misa@fp.ul.pt

## Sumário

*O desenvolvimento de software para crianças requer atenção específica a um conjunto de detalhes que variam desde as metáforas utilizadas, modo de interacção ou até a linguagem utilizada. Estes aspectos ganham ainda mais relevância quando o software a desenvolver está relacionado com actividades críticas como a terapia do medo, o que requer abordagens específicas desde o início do processo de desenho. Este artigo descreve o processo de desenho de um conjunto de soluções para terapia do medo em crianças, através do uso de dispositivos móveis. Demonstramos as técnicas e procedimentos utilizados, assim como os protótipos resultantes. É igualmente apresentada uma avaliação inicial de resultados e planos para trabalho futuro.*

## Palavras-chave

*Design centrado no utilizador; Terapia; Crianças; Dispositivos móveis*

## 1. INTRODUÇÃO

As crianças são constantemente confrontadas com uma grande variedade de situações susceptíveis de provocarem medo durante o seu dia a dia (e.g., *bullying*, escola, lugares escuros). Em certos casos, essas situações e sentimentos, levam a ansiedade, depressão, podendo mesmo afectar gravemente as suas vidas, sendo necessário recorrer à terapia desde idades muito pequenas. Normalmente, a terapia é composta por um estado de diagnóstico inicial cujo objectivo passa pela compreensão da origem do medo, a sua intensidade e o quanto a criança fica afectada no seu dia a dia. Para isso, os terapeutas utilizam artefactos e questionários em papel nas sessões de terapia e, caso necessário, em casa ou até mesmo nas próprias situações responsáveis pela sensação de medo. Contudo, este tipo de procedimentos levanta várias questões, tanto para terapeutas como para pacientes, especialmente em crianças mais novas. Para os terapeutas, torna-se difícil gerir a quantidade de dados recolhidos assim como recolher informação relevante durante as várias situações que podem causar problemas à criança. Para as crianças, a adesão à terapia é reduzida e para os pacientes mais novos (especialmente para os que ainda não sabem ler) a utilização dos questionários tradicionais revela-se, por vezes, problemática.

Recentemente, introduziram-se ferramentas de software para dar suporte à psicoterapia [4][5]. Estas vão desde simples questionários para medir ansiedade [1], até exercícios mais completos de relaxamento e respiração

[3]. Algumas dessas ferramentas suportam a cooperação entre terapeutas e pacientes com algum nível de ajuste, resultando em artefactos que servem contextos específicos [6]. Apesar do uso de software de suporte à terapia não ser novo, é necessário ter em conta a especificidade quando os principais intervenientes são crianças [2]. Algumas das ferramentas existentes, apesar de serem facilmente compreendidas e não obstante os bons resultados demonstrados, são direccionadas a adultos e ou muito específicas a algum tipo de perturbações ou, em alguns casos, demasiado genéricas, não sendo possível adaptar o seu uso por crianças.

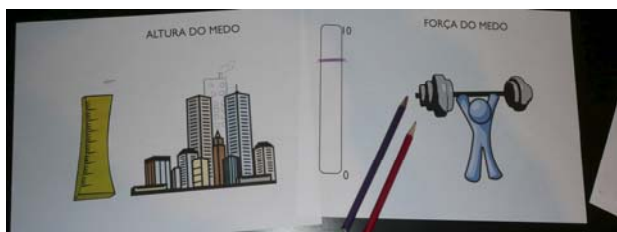
Neste artigo descrevemos o processo inicial de desenvolvimento de um conjunto de protótipos cujo objectivo passa por suportar a terapia do medo em crianças (e.g., entre 5 a 14 anos). O trabalho foi realizado por duas equipas de investigação, uma de engenharia informática e outra de psicologia e psicoterapia, assim como com a colaboração de terapeutas.

## 2. REQUISITOS INICIAIS E PROCESSO DE DESENHO

O processo de desenho iniciou-se com um conjunto de sessões de *brainstorming*. As duas equipas envolvidas discutiram diversas questões enquanto que a equipa de psicoterapia descrevia e explicava os procedimentos utilizados, artefactos e os objectivos de cada terapia. No seguimento destas sessões, e com o objectivo de clarificar certas questões de uma perspectiva diferente,

foram realizadas um conjunto de entrevistas com crianças em terapia, assim como com alguns pais. Estas entrevistas foram realizadas para uma melhor compreensão dos hábitos e à-vontade das crianças para com os computadores e tecnologia, mas também para clarificar como os seus problemas afectam o seu dia a dia, com especial foco em compreender como e quando estas realizam as tarefas relacionadas com a terapia.

Um segundo conjunto de terapeutas providenciou detalhe nos objectivos dos artefactos existentes, nomeadamente no ajuste destes à especificidade dos pacientes ou situações, assim como nos problemas que geralmente afectam os pacientes nas sessões onde são usados. Estes estudos iniciais demonstram que as crianças sentem relutância em usar este tipo de artefactos em papel, o que faz com que os terapeutas tenham de encontrar novas formas de recolha de informação e medição do medo através de desenhos e metáforas.



**FIGURA 1. Artefactos em papel para medir o medo.**

A figura 1, representa um artefacto em papel, onde é pedido ao paciente para registar (e.g., marcar directamente com uma caneta na escala – no lado esquerdo da imagem) a quantidade (ou tamanho) do medo que este sente numa determinada situação. Para este exercício em particular, os terapeutas comentaram que frequentemente os pacientes perguntam se, em vez de marcar na régua, podem desenhar um novo edifício cuja altura represente o tamanho do medo. No lado direito, um segundo artefacto mostra um desenho simples com uma medição do lado esquerdo, no qual o paciente aponta a força do medo. Estes métodos de recolha de informação são geralmente usados no início do processo de terapia, quando os terapeutas tentam utilizar os artefactos e metáforas mais ajustadas para cada criança. Esta abordagem torna-se essencial para cativar o paciente, assim como oferecer-lhe uma abordagem que lhe é fácil de compreender e que, ao mesmo tempo, permite quantificar os seus sentimentos. Contudo, e uma vez que o suporte é em papel, existe um conjunto de questões que afectam estes procedimentos. É comum a uma criança querer editar uma intensidade previamente seleccionada num desenho específico, ou até não entender o que lhe é pedido. Pode-se ainda colocar o problema da criança ter dificuldade em escrever, não conseguindo assim fornecer a informação necessária para alguns desenhos (e.g., escrever o peso do medo). É necessário ainda considerar que o principal desafio em usar este tipo de abordagem, passa por persuadir a criança a registar as situações na qual sentem medo, especialmente quando está longe do terapeuta, situações

nas quais a motivação é mais reduzida, mas os dados são ainda mais importantes de recolher para uso durante o processo de terapia.

Quando questionadas acerca do uso de tecnologia, as crianças responderam muito positivamente. A maioria afirmou usar frequente jogos e dispositivos portáteis (e.g., telefones, leitores de música), sentido-se confortável a usá-los e também extremamente motivada, mesmo aquelas menos experientes no uso de computadores. Os pais, apesar de serem reticentes a deixarem os seus filhos mais novos usar computadores, confirmaram que a maior parte deles permite, e fornece, telemóveis e dispositivos móveis de modo a poderem estar em contacto com eles.

Destes estudos iniciais identificaram-se os seguintes requisitos:

- Permitir que as crianças expressem os seus sentimentos através do uso de metáforas facilmente compreendidas por estas.
- Introduzir soluções tecnológicas de modo a permitir a resposta a alguns dos problemas que afectam o diagnóstico e terapia.
- Utilizar tecnologia adequada à diversidade de situações onde as crianças têm de completar as suas tarefas (e.g., dispositivos móveis).
- Os artefactos têm de ser interactivos e com uma forte componente gráfica, de modo a facilitar a tarefa para crianças com dificuldades em leitura.
- Suportar a interacção directa, assim como a manipulação dos componentes que possibilitam a medição dos sentimentos, sem requerer experiência no uso de computadores, usando modalidades apelativas.
- Reduzir (ou até mesmo eliminar) a introdução de texto para aquelas que não conseguem escrever.
- Enfatizar acções e atitudes positivas através do uso de imagens que sejam ao mesmo tempo esclarecedoras e apelativas, promovendo o uso dos artefactos terapeuticos e adesão à terapia.

### 3. PROTÓTIPOS DE BAIXA FIDELIDADE

Tendo em conta os requisitos iniciais, foram construídos alguns esboços preliminares (figura 2). Estes utilizam dispositivos móveis (e.g., *smart-phones*) e emulam, e melhoram, tirando partido da tecnologia, três artefactos dos estudados, nomeadamente aqueles que medem a intensidade do medo através da sua altura, tamanho e força. Coube aos terapeutas a validação destes protótipos.

Para compreender correctamente o uso desta ferramenta em situações reais, simulou-se um conjunto de sessões de terapia, com o auxílio de três crianças com idades compreendidas entre os 9 e 10 anos. Estes testes realizaram-se em diferentes cenários dentro de uma escola, com o objectivo de compreender o comportamento das crianças ao usar estes protótipos.



**FIGURA 2.** Protótipos de baixa fidelidade iniciais, usados nas sessões de avaliação iniciais. O *designer* muda os vários desenhos enquanto a criança utiliza o dispositivo.

Todas estas sessões foram planeadas meticulosamente e validadas pelos terapeutas. Os objectivos, procedimentos e equipamento usado foram explicados a cada participante. Utilizou-se a técnica *Wizard-of-Oz* para permitir que os participantes interagissem com os vários ecrãs, tendo assim uma simulação do comportamento da ferramenta de acordo com as suas acções. A figura 2 mostra uma criança a interagir com um protótipo de baixa fidelidade, durante uma das sessões preliminares. Cada sessão demorou menos de quinze minutos, seguindo-se uma pequena entrevista.

Apesar dos pacientes mais novos sentirem algumas dificuldades em perceber o conceito por trás da ferramenta, assim como em imaginar as funcionalidades da aplicação através da interacção com esboços em papel, os resultados foram muito positivos. A utilização de técnicas específicas de prototipagem para dispositivos móveis (uso de protótipos físicos com o mesmo aspecto e peso dos dispositivos móveis) [6] foi importante pois forneceu uma experiência mais realista. Na verdade, o interesse e entusiasmo dos pacientes foi maior que o esperado, assim como para com o aspecto e organização visual e com as sequências sugeridas. Aliás, algumas das novas ideias introduzidas no software foram obtidas directamente da observação das expectativas das crianças para com os artefactos em papel. Em particular, o uso de metáforas funcionou não só como meio de medir o medo, mas também como modalidade de interacção, e.g., o facto de pressionarem uma balança como modo de medir o peso do medo (figura 2).

Adicionalmente, ficou claro que que a abordagem mais usada para interagir com os componentes interactivos (explicados no início de cada estudo), foi através da manipulação directa dos mesmos.

#### 4. CATÁLOGO DE ARTEFACTOS DE SOFTWARE

Como seguimento dos resultados positivos dos protótipos iniciais de baixa fidelidade, foi desenvolvido um conjunto de artefactos de software interactivos (figuras 3 e 4). Estes foram baseados nos artefactos em papel, considerando os requisitos iniciais, juntamente com os resultados da avaliação com protótipos de baixa fidelidade.



**FIGURA 3.** Protótipos de software baseados nos protótipos de baixa fidelidade.

#### 5. AVALIAÇÃO E RESULTADOS PRELIMINARES

Para validar o software desenvolvido realizaram-se diversos testes. Estes, foram conduzidos por uma equipa de terapeutas que, durante várias sessões, apresentaram a ferramenta aos seus pacientes para estes a usarem.

As sessões ocorreram no espaço de uma semana no consultório de um dos terapeutas. Durante este período, o software foi utilizado por um total de oito crianças e adolescentes. Todos eles eram estudantes do 4º ao 9º ano de escolaridade, sendo que todos frequentavam terapia para ansiedade e perturbações associadas. Adicionalmente a esta amostra, participaram ainda mais cinco terapeutas, assim como quatro pais. Todos os participantes completaram várias tarefas com os artefactos. Foi-lhes pedido para interagirem várias vezes com todos os ecrãs (e respectivas metáforas e componentes). Posteriormente foi efectuada uma pequena entrevista assim como um questionário a cada um dos participantes. Tanto os testes, como as entrevistas e os questionários, foram conduzidos e controlados pelos terapeutas, com o objectivo de escolherem a melhor aproximação e linguagem, mas também para as crianças não estranharem e se sentirem confortáveis.

O resultado das entrevistas foi claramente positivo. Os vários participantes manifestaram o quanto gostaram das ferramentas e dos diversos componentes, assim como a usabilidade dos mesmos. Três dos oito pacientes deram a nota 7 numa escala de 1 a 7. Apenas um dos participantes, um adolescente, classificou o software como um nota 3 sob o argumento de estes serem muito infantis, enquanto que as restantes quatro crianças deram a nota 6. Em relação à usabilidade, cinco dos oito participantes deram nota máxima, sendo que os restantes três classificaram os protótipos com a nota 6. O questionário efectuada teve como objectivo compreender se a ferramenta desenvolvida tinha um bom nível de usabilidade assim como compreender se esta era preferida aos artefactos em papel. Todos os participantes afirmaram que tinham compreendido o que lhes era pedido em cada ecrã, onde tocar e com que componentes interagir. Adicionalmente, todos concordaram que seriam



capazes de usar a ferramenta, mesmo quando longe do seu terapeuta, sozinhos, sem recorrerem ajuda.

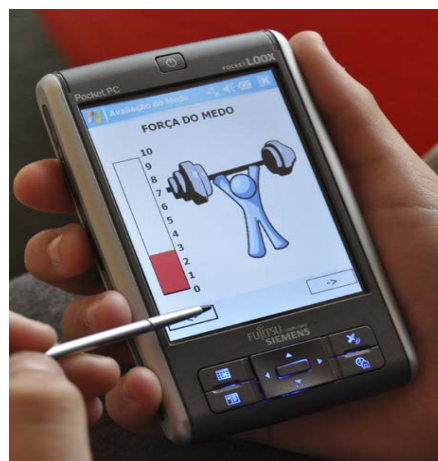
Finalmente, 7 dos 8 participantes acharam que era mais fácil usar os artefactos digitais em detrimento das suas versões em papel. No final do questionário estava uma pergunta de resposta aberta onde era pedido às crianças para escreverem livremente as suas sugestões. Embora os resultados, quer das entrevistas quer dos questionários, tenham sido bastante positivos, todos os participantes tiveram sugestões sobre como melhorar as ferramentas. Os comentários mais comuns entre crianças passavam por ter *feedback* audio e ainda mais *feedback* visual (i.e., por exemplo, imagens de um monstro que aumenta ou diminui de acordo com a quantidade de medo ou a figura humana ganhar mais músculos à medida que o medo é menor). Os pais e terapeutas também se mostraram bastante satisfeitos com as funcionalidades seleccionadas e com a usabilidade e interactividade da ferramenta. Os pais foram unânimes ao constatarem que seria muito mais fácil convencer os seus filhos a realizarem as suas tarefas relacionadas com terapia em casa, utilizando este tipo de aproximação. Quanto aos terapeutas, ficaram extremamente entusiasmados com a possibilidade de poder utilizar dispositivos móveis com os seus pacientes, especialmente com a selecção de ferramentas que podem ser usadas por crianças sem experiência em interacção com computadores.

## 6. DISCUSSÃO E TRABALHO FUTURO

As experiências realizadas durante o período de avaliação demonstraram que as crianças envolvidas responderam de forma bastante positiva às ferramentas apresentadas. Acreditamos que essa resposta é resultado directo dos dispositivos usados e das funcionalidades interactivas, pois estas permitem que as crianças expressem os seus sentimentos de uma maneira atraente, fácil e onde têm maior controlo. Esta crença está em concordância com as opiniões dos terapeutas envolvidos, que constatarem a facilidade de adaptação à tecnologia por parte dos seus pacientes mais novos e o facto destes dispositivos também enfatizarem a sua excitação e vontade a participar nas várias tarefas. Contudo, como ficou explícito nas sugestões e nos resultados globais, é necessário algum cuidado com a selecção do conteúdo apresentado, assim como com o fornecimento de várias alternativas na classificação dos sentimentos. Neste último, a capacidade de interacção com os desenhos revelou-se crucial, cativando o interesse da criança em usar as ferramentas.

Tanto os pais como os terapeutas foram peremptórios em constatar que a ferramenta pode melhorar o compromisso das crianças com a terapia. Contudo, será certamente necessário considerar a constante evolução e inclusão de novas funcionalidades na aplicação, no sentido de cativar o interesse da criança durante longos períodos de tempo, permitindo também a utilização de aproximações diferentes para pacientes com problemas distintos. Mesmo assim, quando comparados com os artefactos

tradicionais, em papel, estes novos artefactos digitais ultrapassam as limitações iniciais que motivaram este trabalho: (1) facilidade no processo de edição; (2) fácil interacção para crianças de idade reduzida (incluindo crianças que ainda não sabem ler), assim como a gestão de informação para os terapeutas; (3) e finalmente, o aumento do entusiasmo e o compromisso para com a terapia.



**FIGURA 4. Criança a testar o protótipo num PDA. O ecrã permite que esta quantifique o medo indicando a sua força, através da escala do lado esquerdo.**

Como trabalho futuro, está previsto o refinamento destes artefactos, bem como o desenvolvimento e a integração de outros. Estão igualmente planeados a realização de novos testes, mais prolongados e com um grupo de utilizadores mais alargado.

## 7. REFERÊNCIAS

- [1] Grasso, M. A., Clinical Applications of Hand Held Computing. 17th IEEE Symposium on Computer Based Medical Systems, pp.141-146, 2004.
- [2] Hailpern, J., Encouraging Speech and Vocalization in Children with Autistic Spectrum Disorder, SIGAccess Newsletter, pp.47-52, 89, 2007.
- [3] Newman M. G., 2004. Technology in psychotherapy: an introduction. In Journal of Clinical Psychology. No. 60 (2) pp 141-145
- [4] Proudfoot, J., 2004. Computer-based treatment for anxiety and depression. Neuroscience and Biobehavioral Reviews, No. 28 pp 353-363.
- [5] Przeworski, A., Newman, M., 2004. Palmtop computer-assisted group therapy for social phobia. Journal of Clinical Psychology. 60(2)179-188.
- [6] Sá, M., Carriço, L., Antunes, P., Ubiquitous Psychotherapy, Pervasive Computing, Special, vol. 6, nº 1, pp. 20-27, 2007, IEEE.



## **Sessão VIII**

### **Interacção em Contextos Diversos**

Aveiro, 13, 14 e 15 de Outubro de 2010

# Estudo Comparativo de Aplicações para a Construção de Modelos LEGO<sup>TM</sup>

Daniel Mendes   Alfredo Ferreira  
Instituto Superior Técnico / INESC-ID  
R. Alves Redol, 9, 1000-029 Lisboa, Portugal

danielmendes@ist.utl.pt, alfredo.ferreira@inesc-id.pt

## Resumo

*A construção de modelos LEGO<sup>TM</sup> é um passatempo popular não apenas para crianças e jovens, mas também para adultos de todas as idades. Acompanhando a evolução tecnológica, existem actualmente diversas aplicações que permitem a construção de modelos virtuais. Neste documento é descrito um teste com utilizadores realizado com o objectivo de comparar três destas aplicações. Com base nos resultados obtidos, tentou-se compreender quais as abordagens que mais agradam e desagradam aos utilizadores no que se refere à interacção com este tipo de aplicações. Pretende-se assim reunir informação que suporte o desenvolvimento de aplicações para construção de LEGO<sup>TM</sup> virtual baseadas em novos paradigmas de interacção.*

## Palavras-Chave

*Estudo Comparativo, Testes com Utilizadores, Ferramentas de Modelação, LEGO<sup>TM</sup>*

## 1 INTRODUÇÃO

Ao longo da sua infância muitas pessoas brincaram com blocos de construção que encaixavam uns nos outros de modo a criar aquilo que a imaginação ditava. LEGO<sup>TM</sup> é o mais famoso fabricante deste tipo de brinquedos, globalmente conhecido pelos seus blocos de construção e não só. Na realidade, construir modelos LEGO<sup>TM</sup> é uma actividade partilhada por pessoas de todas as idades. Como complemento aos tradicionais blocos de plástico, existem actualmente diversas aplicações que permitem construir modelos de forma virtual.

A construção de modelos LEGO<sup>TM</sup> virtuais permite uma maior liberdade criativa, visto não haver a necessidade de adquirir os ditos blocos, tendo-se disponível uma grande variedade destes. A popularidade destas aplicações tem vindo a aumentar e, inclusivamente, já foram realizados estudos comparativos entre a construção de modelos de LEGO<sup>TM</sup> com peças físicas e a sua criação com ferramentas virtuais [Baradaran 06]. Uma das conclusões deste estudo foi o facto de a construção no ambiente real ser mais rápida, para utilizadores novíços, que em ambiente virtual. Não encontramos, no entanto, nenhum focado em comparar diferentes programas e abordagens para a construção de LEGO<sup>TM</sup> virtual.

Neste artigo apresenta-se um estudo comparativo de aplicações para a criação de modelos LEGO<sup>TM</sup> virtuais. Este estudo envolveu um conjunto de utilizadores novíços e surge no âmbito do projecto LTouchIT. Este projecto tem

como objectivo final o desenvolvimento de uma aplicação de criação de modelos LEGO<sup>TM</sup> para uma superfície multi-toque. Esta aplicação deverá oferecer aos utilizadores interacção mais natural do que as tradicionais ferramentas WIMP e mais próxima do modelo mental que estes já possuem sobre a manipulação das peças no mundo real.

Com esta análise pretende-se compreender do ponto de vista dos utilizadores, qual a melhor abordagem a seguir no que toca à construção de modelos LEGO<sup>TM</sup> virtuais. Nomeadamente, serão identificados os pontos fortes e fracos das diversas aplicações, assim como aspectos que poderiam ser melhorados. A partir da análise dos dados obtidos será possível encaminhar o desenvolvimento de uma aplicação para uma superfície multi-toque para a criação deste tipo de modelos, orientada para as necessidades e expectativas dos utilizadores.

No restante do documento iremos apresentar as três aplicações incluídas neste estudo e a metodologia utilizada para a realização dos testes. Iremos ainda descrever as condições em que decorreu a experiência e qual o perfil dos utilizadores participantes. Por fim, apresentaremos os resultados obtidos e as respectivas conclusões, indicando o caminho para trabalho futuro.

## 2 SISTEMAS A TESTAR

Para efectuar o referido estudo, foram efectuados testes com utilizadores, nos quais foram consideradas três aplicações, as vulgarmente utilizadas para esta finalidade: LEGO<sup>TM</sup> Digital Designer, MLCad e LeoCAD.

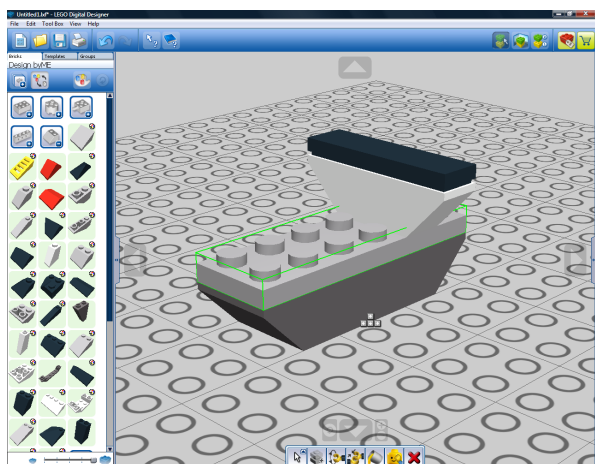


Figura 1: Ecrã da aplicação LEGO™ Digital Designer.

De seguida encontram-se apontadas as principais características de cada uma destas aplicações, nomeadamente no que se refere à procura, selecção, manipulação das peças e controle da câmara.

## 2.1 LEGO™ Digital Designer

O primeiro destes sistemas, o LEGO™ Digital Designer (LDD) [ldd10], é uma aplicação proprietária da LEGO™, pelo que é uma aplicação fechada. A modelação é feita num ambiente tridimensional e possui um sistema de ajuda eficiente nas ligações entre as peças, sendo a única aplicação neste estudo que impede a sobreposição de peças.

A aplicação apresenta as peças listadas pelo seu *thumbnail*, sendo possível incluir nesta lista peças iguais de cores diferentes, ou apenas uma peça representativa e escolher posteriormente a respectiva cor. Existe uma grelha sempre visível, simulando uma placa de construção, conforme se pode ver na Figura 1. A movimentação das peças faz-se exclusivamente no plano desta grelha, à qual se adaptam, ligando-se a outras peças que já estejam no modelo sempre que o programa considere conveniente. A rotação das peças faz-se apenas segundo dois eixos, utilizando-se as setas direccionais para cada rotação de 90°, não sendo possível rotações intermédias.

O movimento da câmara é efectuado arrastando-se o rato e para aproximar e afastar recorre-se à roda de deslocamento (*scroll wheel*). Para arrastar a câmara o sistema mapeia o movimento do rato, não seguindo o cursor de forma exacta. No que toca à rotação da câmara, esta roda em torno dos eixos horizontal e vertical da vista actual do modelo, orbitando em torno de um ponto e sem nunca se inclinar.

## 2.2 MLCad

O MLCad [mlc10] é um sistema de CAD aplicado à construção de modelos de LEGO™ e é principalmente orientado à criação de instruções de construção. À semelhança das tradicionais ferramentas de modelação 3D,

o MLCad possui quatro *viewports*, conforme se pode ver na Figura 2. Estes representam diferentes vistas do modelo: três ortogonais (topo, frente e lateral) e uma perspectiva. No entanto, apenas é possível realizar alterações ao modelo nas vistas ortogonais, sendo a vista perspectiva apenas para visualização. Utiliza como base a biblioteca de peças LDraw [ldr10].

A pesquisa das peças é efectuada segundo agrupamentos das peças, os quais são identificados textualmente, e as respectivas peças encontram-se listadas em dois locais: numa lista textual e noutra pelos seus *thumbnails*. O deslocamento das peças efectua-se apenas nos planos das vistas ortogonais, os quais não possuem qualquer tipo de grelha auxiliar, não existindo qualquer restrição na posição da peça nem ligações com outras peças. Existem seis botões na interface para controlar a rotação das peças, dois para cada um dos três eixos, para rotações de 90° em sentidos opostos.

A movimentação da câmara é feita arrastando um ponto da vista com o rato. A roda de deslocamento é utilizada para aproximar e afastar. A aplicação proporciona rotações da câmara em torno dos três eixos, sendo o modelo centrado após cada rotação.

## 2.3 LeoCAD

O LeoCAD [leo10], tal como a aplicação anterior utiliza a biblioteca LDraw. As suas opções disponibilizam uma grelha, à qual as peças se adaptam a cada meia unidade, e diversos conjuntos de *viewports*. Tanto a grelha como os *viewports* não vêm activos por omissão, mas podem ser activados se necessário. Para alternar entre as diversas funções, como movimentar as peças ou rodar a câmara, é sempre necessário recorrer aos botões da interface.

As peças são apresentadas através de uma lista com grupos, cuja identificação, tanto das peças como dos grupos, é apenas textual. Existe ainda uma janela com a pré-visualização da peça seleccionada. Durante a sua manipulação, as peças deslocam-se normalmente num plano horizontal, sendo que para as movimentar num plano

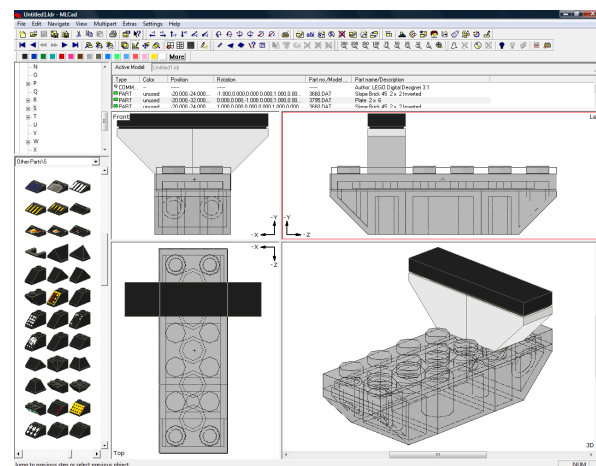


Figura 2: Ecrã da aplicação MLCad.

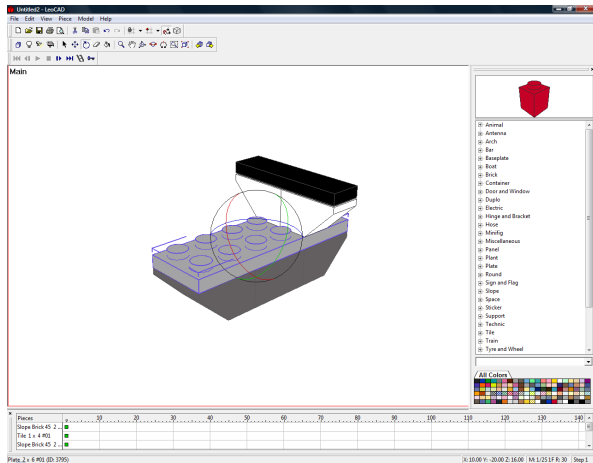


Figura 3: Ecrã da aplicação LeoCAD.

vertical é necessário premir o botão do lado direito do rato ou seleccionar o eixo correspondente. No caso de se utilizar *viewports* ortogonais, as peças movimentam-se paralelamente aos respectivos planos de visualização. O movimento do rato é mapeado segundo os eixos, não existindo um ponto que siga o cursor.

Para a rotação das peças foi seguida uma abordagem idêntica à proposta por Ken Shoemake [Shoemake 92] para especificar orientação tridimensional usando um rato tradicional. A implementação usada no LeoCAD recorre a uma esfera, visível na Figura 3, que roda a peça em intervalos de 30° relativamente a cada um dos três dos eixos. Efectuando a rotação clicando numa das circunferências faz com que a rotação se faça exclusivamente segundo um eixo.

O arrastar da câmara para a mover é feito de forma muito sensível ao movimentar o rato. Pelo contrário, a roda de deslocamento utilizada para afastar e aproximar a câmara é pouco sensível. De sublinhar que esta aplicação oferece diversas formas de rotação da câmara: rodar em torno dela própria, orbitar à volta de um ponto e inclinar.

### 3 METODOLOGIA DOS TESTES

Com o objectivo de determinar os principais aspectos positivos e negativos de várias aplicações descritas na secção anterior, foram conduzidos um conjunto de testes com utilizadores, seguindo a metodologia sugerida por Preece, Rogers and Sharp [Preece 02].

A sessão de testes com os utilizadores estava organizada em três etapas, identificadas na Tabela 1, juntamente com a respectiva duração. Segue-se uma explicação de cada uma das referidas etapas.

#### 3.1 Introdução à experiência e às aplicações

No início da experiência começou-se por explicar ao utilizador o propósito da sessão e tentou-se familiarizá-lo com as aplicações em causa. Para o efeito, foram apresentadas as principais funcionalidades de cada aplicação, tendo-se seguido uma fase de treino, composta por um período de três minutos para interagir livremente com cada

uma das aplicações.

#### 3.2 Experiência de utilização das aplicações

Após a fase de treino, foi solicitado a cada utilizador que completasse uma tarefa em cada uma das aplicações. A tarefa consistia na construção de um modelo simples, presente na Figura 4, que envolvia procura, selecção, manipulação, rotação e colocação das peças. Para a execução da tarefa em cada aplicação foi disponibilizado um tempo máximo de cinco minutos. Após este tempo seria solicitado ao utilizador que passasse para a próxima aplicação. A ordem pela qual os utilizadores utilizaram as aplicações para efectuarem a tarefa foi aleatória para cada um.

Nesta etapa, foi pedido aos utilizadores que partilhassem os seus pensamentos com o observador enquanto realizavam a tarefa. Desta forma conseguimos compreender melhor quais as principais dificuldades com que os utilizadores se deparavam durante a realização da tarefa. Uma vez que esta abordagem é incompatível com medições de desempenho, estas foram adiadas para uma futura ronda de testes com utilizadores.

Etapa	Nome	Duração
1	Introdução à experiência	10 min
2	Experiência de utilização	15 min
3	Preenchimento do questionário	5 min

Tabela 1: Etapas dos testes com utilizadores.

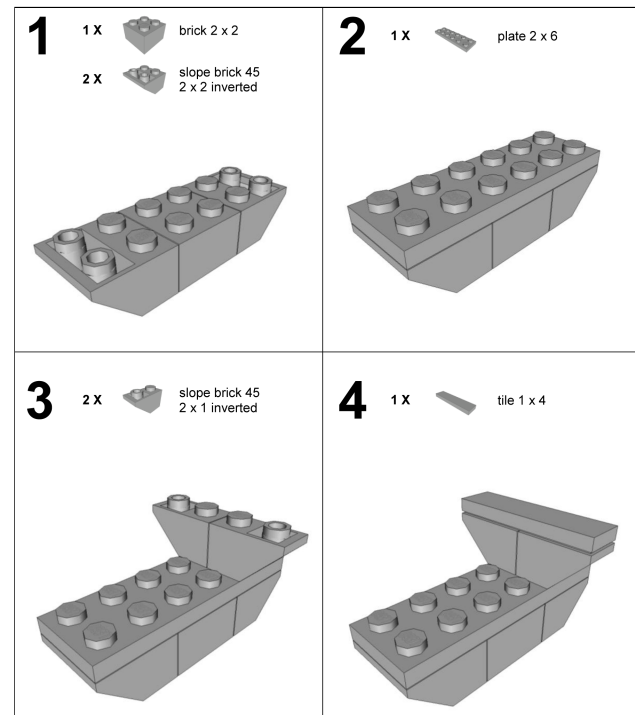


Figura 4: Tarefa a realizar pelos utilizadores.

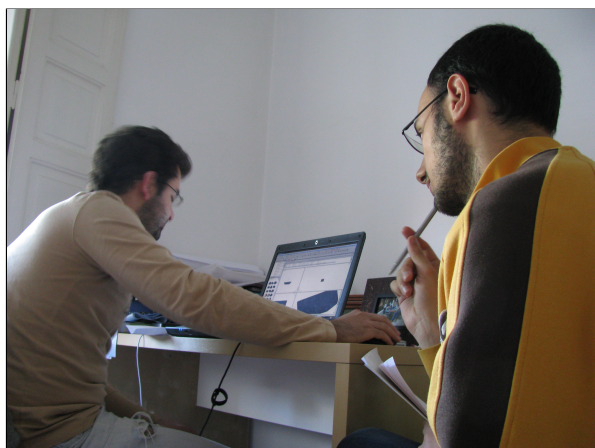


Figura 5: Realização de uma sessão de testes com um utilizador.

### 3.3 Preenchimento do questionário

Após ter realizado a tarefa nas três aplicações aleatoriamente, foi pedido a cada utilizador que preenchesse um curto questionário. Este questionário focava a experiência com cada uma das aplicações. Esperava-se assim identificar quais os pontos fortes e fracos em cada uma das aplicações, com especial relevo para a procura e selecção de peças, assim como a sua manipulação e a da câmara.

O questionário era composto por um primeiro conjunto de questões de escolha múltipla para identificar o perfil do utilizador, seguido por cinco questões de classificação dos vários aspectos de cada aplicação, expostas na Tabela 2, cuja resposta consistia em avaliar cada aspecto segundo uma escala de Likert com 4 valores, sendo 1 uma má classificação e 4 uma classificação positiva. Por fim, existia uma pergunta de resposta aberta, na qual era solicitado aos utilizadores que apontassem aspectos relevantes, quer positivos quer negativos, das várias aplicações.

## 4 TESTES COM UTILIZADORES

Os testes envolveram 21 utilizadores e foram realizados num ambiente calmo e controlado, proporcionando uma execução dos testes sem interrupções, ilustrado na Figura 5. A tarefa foi executada num computador portátil, equipado com ecrã de 15.6" na resolução de 1280x800 pixels e sobre o sistema operativo Microsoft Windows Vista. Foi utilizado um rato externo sem fios, de modo a proporcionar uma interacção mais familiar a todos os utilizadores, em vez do touchpad incorporado.

As idades dos utilizadores foram variadas, sendo que o utilizador mais novo tinha 9 anos. A grande maioria (76%) dos utilizadores situava-se na faixa etária entre os 18 e os 34 anos e mais de metade (62%) dos utilizadores detinham um curso superior.

Todos os utilizadores possuíam alguma experiência na utilização de computadores, sendo que 48% asseguraram ser muito experientes. No que toca à utilização de programas de manipulação de imagens e de modelação tridimen-

sional, a maior parte dos utilizadores afirmaram ter, pelo menos, alguma experiência (76% e 71%, respectivamente).

Dois terços dos utilizadores constataram terem feito construções em LEGO<sup>TM</sup> tradicional muitas vezes, sendo que o terço restante fez algumas vezes. No entanto, nenhum dos utilizadores tinha utilizado previamente um programa de construção de modelos LEGO<sup>TM</sup> virtuais.

## 5 ANÁLISE DOS RESULTADOS

As classificações resultantes das avaliações efectuadas pelos utilizadores, relativas a cada questão para cada aplicação, estão representadas através da sua média no gráfico da Figura 6. Utilizou-se o teste de Shapiro-Wilk e constatou-se que os valores atribuídos pelos utilizadores seguem uma distribuição normal para  $\alpha = .05$  ( $.000 \leq \text{sig.} \leq .012$ ). Estes valores foram depois analisadas usando uma One-Way ANOVA. Para as questões 1,3,4 e 5, as médias obtidas são significativamente diferentes para qualquer valor de  $\alpha$  ( $\text{sig.} = .000$ ). Para a questão 2, as médias obtidas são significativamente diferentes para  $\alpha = .01$  ( $\text{sig.} = .002$ ).

Analisando o gráfico da figura 6, pode-se concluir que o LDD é o que melhor cumpre a função de construção de modelos LEGO<sup>TM</sup> nos vários aspectos estudados, garantindo ao utilizador uma maior satisfação na sua utilização. Considerando a primeira questão no gráfico, constata-se que a distância que separa o LeoCAD do MLCad é bastante superior à que separa este último do LDD. Tal pode ser justificado, para além de outros aspectos mencionados nas secções seguintes, com a necessidade de recorrer constantemente a botões na interface para alternar entre funções no MLCad, o que causou, por várias vezes, confusão aos utilizadores.

Embora a utilização de medidas de desempenho seja incompatível com a técnica *think aloud* utilizada, os tempos obtidos pelos utilizadores na realização das tarefas e o número de utilizadores que completaram as tarefas com sucesso no tempo limite de cinco minutos encontram-se nas Figuras 7 e 8 respectivamente, a título ilustrativo. Apesar de o objectivo deste estudo recair sobre as preferências dos utilizadores relativamente a diferentes abordagens para a criação de modelos LEGO<sup>TM</sup> virtual, ao invés de analisar o desempenho nas diferentes aplicações, nota-se que, mais uma vez, a da aplicação da LEGO<sup>TM</sup> apresenta melhores resultados.

De seguida são analisados os principais aspectos relativos à modelação de LEGO<sup>TM</sup> virtual, nomeadamente a selecção e manipulação das peças e controlo da câmara, tendo como base as respostas dos utilizadores ao questionário, assim como os seus comentários aquando da experiência de utilização.

### 5.1 Procura e Selecção de Peças

No que toca à procura e selecção de peças, as opiniões dos utilizadores, manifestadas na segunda questão, foram um pouco variadas, tendo sido, na sua maioria, preferencial em relação ao LEGO<sup>TM</sup> Digital Designer. Embora



Nº	Questão
Q1	De uma forma geral, como classifica a sua satisfação na utilização da aplicação?
Q2	Como classifica a procura e selecção de peças?
Q3	Como classifica a movimentação e colocação das peças no modelo (excluindo a rotação)?
Q4	Como classifica a rotação das peças?
Q5	Como classifica a visualização do modelo e a navegação pela cena?

Tabela 2: Questões presentes no questionário de satisfação.

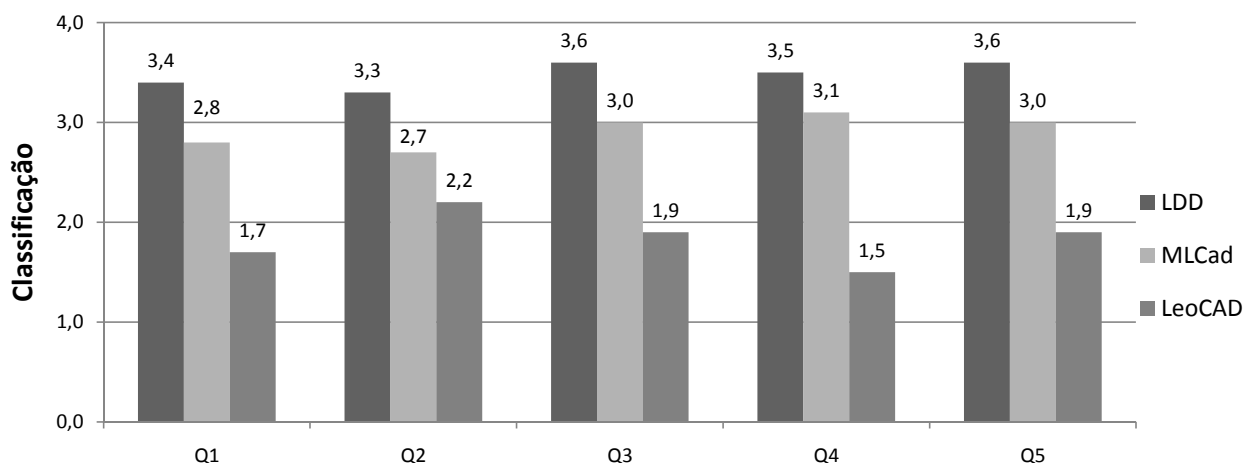


Figura 6: Classificações médias das aplicações nas cinco questões.

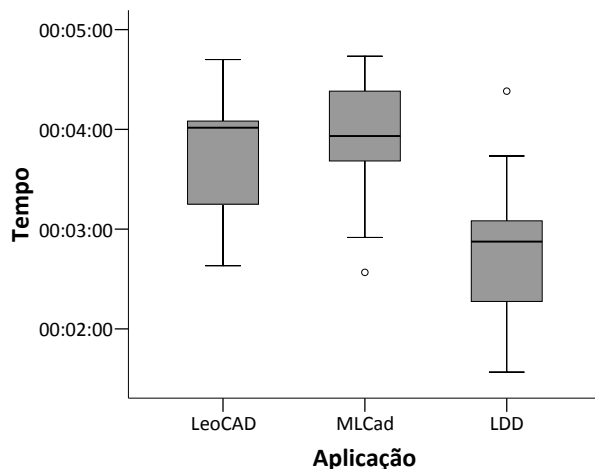


Figura 7: Tempos obtidos na realização das tarefas nas diferentes aplicações.

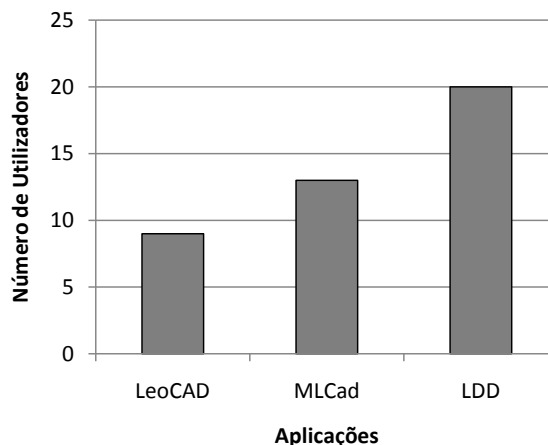


Figura 8: Número de utilizadores que completaram a tarefa com sucesso em cada uma das aplicações.

os utilizadores concordassem que a organização das peças no LDD é a que transmite a sensação de LEGO™ tradicional, o facto de não ser de fácil percepção o grupo a que uma dada peça pertence e, ainda, a tarefa possuir nas suas instruções o nome das peças, fez com que vários utilizadores considerassem a pesquisa por nomes mais fácil. Uma vez que, no LDD, pousando o cursor sobre uma peça, na lista de peças, mostra o seu nome, que geralmente inclui as dimensões da peça, vários utilizadores recorreram a isto

para terem conhecimento destas dimensões, dado que nem sempre foi claro observando apenas o *thumbnail*.

O que permitiu que o MLCad se distanciasse do LeoCAD neste campo, visto que ambos possuem as peças listadas textualmente, deveu-se à primeira também possuir uma listagem visual do conteúdo dos grupos. Não obstante de, no MLCad, nem sempre o *thumbnail* das peças coincidir com a peça seleccionada e algumas peças estarem agru-

padas num grupo específico e outras estarem por ordem alfabética. A possibilidade de arrastar a peça quer pelo nome, quer da lista com imagens foi um aspecto positivo desta aplicação, agradando a utilizadores com diferentes preferências. Houve alguns utilizadores que tentaram arrastar a peça também pelo seu *thumbnail* no LeoCAD, mas sem sucesso. O facto de no LDD não se arrastar a peça para a cena, mas sendo necessário clicar nela para fazer com que ela passe a seguir o cursor e depois clicar quando estiver no sítio pretendido originou alguma confusão e frustração.

De realçar o facto de vários utilizadores terem recorrido, por iniciativa própria, à funcionalidade de copiar e colar que os programas ofereciam, para duplicar algumas peças. Tal possibilidade foi do agrado destes utilizadores, sendo-lhes poupado o esforço de ir à lista de peças e colocar a peça na cena.

## 5.2 Translação e Rotação das Peças

Na terceira questão, relativa à movimentação e colocação das peças no modelo, mais uma vez o LDD foi a aplicação predilecta dos utilizadores. Para isto ajudou possuir um sistema de adaptação das peças umas às outras e à grelha que serve de base à construção. Este sistema, no entanto, originou que, por diversas ocasiões, o programa considerasse uma posição para a peça diferente da que o utilizador pretendia. Por exemplo, o utilizador pretendia colocar uma peça adjacente a outra e o programa colocava-a por baixo. Esta desvantagem foi facilmente controlável movimentando-se a câmara para um ângulo que tornasse as intenções do utilizador mais claras. A ausência de um mecanismo deste género foi apontada como uma falha nos outros dois programas por vários utilizadores, sendo que, por vezes, recorreram às setas direccionais para conseguir posicionar a peça com uma melhor precisão.

Por seu lado, o LeoCAD não possui um mecanismo de colocação das peças. No entanto, este programa tenta adaptar as peças a outras que já existam na cena, mas apenas quando se arrasta a peça da lista e nunca tendo ficado na posição que os utilizadores pretendiam. Uma vez estando a peça colocada na cena, quando o utilizador a move este mecanismo deixa de existir. Este semi-sistema de adaptação das peças foi mais uma fonte de frustração para os utilizadores. Adicionalmente, o facto das peças apenas se movimentarem num plano horizontal, sendo necessário uma acção especial para a mover na vertical e não seguirem exactamente o cursor enquanto são arrastadas, confundiu vários utilizadores.

A rotação das peças é claramente o aspecto mais fraco do LeoCAD, causando uma certa frustração a vários utilizadores. Embora seja a que fornece uma maior liberdade na rotação, esta liberdade não é do agrado dos utilizadores, tendo originado várias rotações que não as pretendidas. Totalmente oposto ao LeoCAD, o sistema de rotação de peças do LDD, só permitindo a rotação segundo dois eixos e ângulos estabelecidos três vezes superiores aos da primeira aplicação, foi o que mais cativou os utilizadores, embora tivesse causado alguma confusão o facto de seleccionar uma peça não ser o suficiente para a poder

rodar, sendo necessário primeiro movimentá-la. Em todas as aplicações houve, por vezes, dificuldade em definir o eixo sobre o qual se deveria efectuar a rotação para que a peça ficasse com a orientação desejada.

## 5.3 Manipulação da Câmara

Por fim, os resultados da última questão, respeitante à manipulação da câmara, revelaram que tanto o LDD como o MLCad foram do agrado dos utilizadores, com uma maior inclinação para o LDD. Em relação ao LDD, tal pode ser explicado devido ao facto do programa dar poucas opções para a rotação da câmara, em comparação com o LeoCAD, o que, mais uma vez, torna a interacção mais simples. Alguns utilizadores chegaram mesmo a evitar efectuar rotações e translações da câmara no LeoCAD, dada a sua complexidade. A elevada classificação dada ao MLCad neste campo é justificável através das diferentes vistas que este possui, evitando a necessidade dos utilizadores terem de controlar a câmara.

A utilização de *viewports* neste último desagradou a quase tantos utilizadores quanto agradou a outros. Esta diferença deve-se, na sua maioria, à experiência dos utilizadores na manipulação de sistemas de CAD ou outras aplicações que possuam um conjunto de vistas semelhante ter ajudado na utilização deste programa. Mesmo para estes utilizadores, foi consensual que as várias perspectivas são práticas, mas não fornecem uma interacção natural, nem o *look-and-feel* de construções LEGO<sup>TM</sup>, havendo aqui vários desafios por resolver.

## 6 CONCLUSÕES

Neste documento realizou-se um estudo comparativo a um conjunto de três programas de construção de modelos LEGO<sup>TM</sup>. Para o efeito, foram efectuados testes com utilizadores que envolveram vinte e um indivíduos. Foi pedido a cada um destes indivíduos que realizasse uma tarefa de construção de um modelo LEGO<sup>TM</sup> nas três aplicações, após a qual lhe foi solicitado que respondesse a um questionário de satisfação. Através da análise dos resultados obtidos, estes testes ajudaram a determinar factores importantes a ter em conta no desenvolvimento de uma aplicação de simulação de blocos de construção. Vários destes factores poderão ser também válidos noutro tipo de aplicações de modelação tridimensional.

Foi frequentemente referido pelos utilizadores que o LDD é a aplicação é mais familiar e a que melhor passa a sensação de construções em LEGO<sup>TM</sup> tradicional. Os utilizadores consideraram mais natural ver as peças listadas pelas suas imagens, mas afirmam ser útil fornecer alguma informação textual como, por exemplo, as suas dimensões. Houve ainda um utilizador que sugeriu que seria interessante possuir um campo de texto para efectuar pesquisas de peças.

Existir um sistema de adaptação das peças e que evite colisões entre elas é importante, contribuindo para uma interacção mais familiar com a aplicação. Ainda no que respeita à simulação do comportamento das peças num cenário de LEGO<sup>TM</sup> físico, um dos utilizadores mais novos

estranhou ser possível retirar uma peça de baixo de outra, sem que a de cima viesse para baixo, como que por efeito da gravidade.

Mostrou-se que os utilizadores preferem dispor de funções que lhes permitam fazer cumprir os objectivos que pretendem de forma mais simples e clara, mesmo que isso implique uma interacção com um pouco menos liberdade. Tal manifestou-se na movimentação e rotação das peças, assim como na manipulação da câmara.

Com base na informação recolhida neste estudo comparativo, esperamos possibilitar o desenvolvimento de novas aplicações para construção de modelos LEGO<sup>TM</sup> virtuais com interfaces mais naturais que as existentes. Mais especificamente, esperamos contribuir para o desenvolvimento de soluções que tirem partido de novos dispositivos de interacção, tais como as superfícies multi-toque.

## 7 Agradecimentos

Este trabalho foi financiado em parte pela Fundação para a Ciência e Tecnologia através do projecto 3DORuS (PTDC/EIA-EIA/102930/2008).

## Referências

[Baradaran 06] H. Baradaran e W. Stuerzlinger. A Comparison of Real and Virtual 3D Construction Tools with Novice Users. *CGVR'06*, 2006.

[ldd10] Lego digital designer. <http://ldd.lego.com/>, visitado em 28 Maio 2010.

[ldr10] Ldraw. <http://www.ldraw.org/>, visitado em 28 Maio 2010.

[leo10] Leocad. <http://www.leocad.org/>, visitado em 28 Maio 2010.

[mlc10] Mlcad. <http://www.lm-software.com/mlcad/>, visitado em 28 Maio 2010.

[Preece 02] Jenny Preece, Yvonne Rogers, e Helen Sharp. *Interaction Design: Beyond Human-Computer Interaction*, capítulo 14. John Wiley and Sons Ltd, 2002.

[Shoemake 92] Ken Shoemake. Arcball: a user interface for specifying three-dimensional orientation using a mouse. Em *Proceedings of the conference on Graphics interface '92*, páginas 151–156, San Francisco, CA, USA, 1992. Morgan Kaufmann Publishers Inc.

Aveiro, 13, 14 e 15 de Outubro de 2010

# Empirical Study of User Behavior on Home Page Layouts: SAPO Portal Case Study

Samuel Almeida  
Dep. Comunicação e Arte  
Universidade de Aveiro  
Aveiro  
sja@ua.pt

Óscar Mealha  
Dep. Comunicação e Arte  
Universidade de Aveiro  
Aveiro  
oem@ua.pt

Ana Veloso  
Dep. Comunicação e Arte  
Universidade de Aveiro  
Aveiro  
aiv@ua.pt

André Luís  
Portugal Telecom / SAPO  
Lisboa  
andre-filipe-  
luis@telecom.pt

---

## Abstract

*The use of eye tracking in usability research has increased in the past years. The work presented in this paper reports on an empirical study that took place at the University of Aveiro. 58 participants were asked to read news on the SAPO homepage while their eye movements and mouse interaction were recorded. An analysis of their visual and interactive behavior on the news areas of interest was made as well as other areas of the homepage. Acquired results were considered, and reinforced interaction design criteria used in the reconceptualization of the actual SAPO homepage.*

## Keywords

*Eye tracking, visual behavior, interaction design, usability*

---

## 1. INTRODUCTION

The use of eye tracking in usability research and analysis is growing. Some of the first signs of eye tracking date back to 1879, when Louis Émile Javal observed a specific pattern in the reading process (Leggett, January 19, 2010). Over the following years, eye tracking has been employed in several other areas of research: advertising, video games, design and the web. Many of the R&D projects mentioned in the following sections were feasible because eye tracking technology has become economically accessible and technically evolved into a very accurate and non invasive measuring instrument (Duchowski, 2007; Jacob & Karn, 2003).

The work reported in this paper shows how an empirical study that gathered qualitative information, visual attention data and mouse clicks detected user interaction behavior and confirmed reconceptualization criteria of the Portuguese Telecom SAPO search engine portal.

## 2. RELATED WORK

(Mosconi, Porta, & Ravarelli, 2008) analyzed on-line newspapers, focusing mainly on their multimedia content. Using eye tracking technology, their study led them to present several suggestions for on-line newspapers, namely for “news formats”, “links to multimedia sections”, “slideshows”, “ad banners” and “videos”.

(Holmqvist, Holsanova, Barthelson, & Lundqvist, 2003) discuss the differences between reading traditional newspapers and on-line newspapers (net papers), presenting results that indicate that net paper readers scan more but read less than newspaper readers. (Barthelson, 2002) demonstrated through an eye tracking study that visualizing a webpage follows several criteria, including a page's inherent prioritized layout, that content located in the middle of the page receives more attention and that personal interests and intentions also influence reading choices.

(Hsieh-Yee, 2001) analyzed several studies focused on web search behavior that considered both children and adult participants. Based on research, (Hsieh-Yee, 2001) also indicated, that several factors may influence how users search for information. These factors may be related to a user's background and experience with computers, the web and other information retrieval tools. Additionally, information need, domain knowledge, cognitive abilities, affective states, demographics, and the nature of the information also contribute to the way in which the searcher seeks information (Hsieh-Yee, 2001, p. 169). Finally, the type of search task may also affect information search. People use different search strategies depending if the task is fact-based, open ended or object oriented. (Pan, et al., 2004) indicated that the gender of the subjects, the order in which web pages are visualized

and the interaction between site types influence the behavior or web page viewing. These results confirmed work previously done by (Rayner, Li, Williams, Cave, & Well, 2007) that also demonstrated the effects of individual characteristics and stimuli on eye movement behavior.

Other relevant studies are those elaborated by the Poynter Institute<sup>1</sup>, some dating back to the early nineties. The first Poynter study was presented in 1991. Using 90 participants, three newspapers were examined by readers. Results indicated that photos and color are important, and that readers have a common pattern of navigation: readers begin on photos, jump to headlines, cutlines and then text (Garcia, 1991). Another study took place between 1998 and 2000 in collaboration with Stanford University (Lewenstein, Edwards, Tatar, & DeVigal, 2000). The study analyzed on-line newspapers and 67 testers were asked to read different newspapers based on preferences and usual habits. Contrary to what was seen in the previous study related to printed news; in on-line newspapers, the main entry point is through text, suggesting that readers prefer pictures to be separated from the news piece. A third study from the Institute titled “Eye Track III” (Outing & Ruel, 2004) took place in 2004 and analyzed once again online readers. The 46 participants that took part in the study were asked to read mock news websites as well as multimedia content while their eye movements were registered. The study resulted in several findings: readers’ eyes fix first in the upper left of a page (normally where a logo is placed), then go from left to right; larger headlines (especially located in the left) draw more attention than smaller headlines; text is preferably an entry point into web pages; short paragraphs received twice as much attention as long; ads in the top and left of a page receive the most attention.

(Nielsen & Pernice, 2009) also developed an extensive eye tracking study, having focused on analyzing visualization patterns on websites with a business goal. Over 300 subjects, aged 18 to 64, participated in a study that collected data related to the way people visualize websites, navigation menus, images and advertisements.

### 3. OBJECTIVES

The main goal of this study was to gather information able of revealing relevant interaction behavior of SAPO users while finding and reading news exclusively on the SAPO homepage. Reconceptualization of SAPO’s homepage articulated with this objective and was partially driven as a consequence of the results reported in this paper.

### 4. METHODS & TECHNIQUES

The present study took place at the University of Aveiro, Portugal and all related data was gathered during one week in March, 2009. The following sections describe details of what can be classified as an empirical study based on eye tracking technology and intrinsic registration of user interaction behavior.

#### 4.1 Sample description

The sample of the study consisted in 58 participants which were selected randomly upon volunteering. Partic-

ipants’ ages ranged from 18 to 43 and were all members of the University of Aveiro community. 34 participants were male and 24 were female. The average age of the participants was 24.5. For result analysis purposes, the study sample was divided into 4 groups: participants under the age of 20; participants aged 20-29; 30-39 and participants with 40 or more years of age. Table 1 characterizes the number of participants per age group and the respective percentage in terms of the global number. Because most of the participants were students, the 20-29 age group had the greater portion of participants.

	<20	20-29	30-39	40+
No. P.	9	43	4	2
%	15.5%	74,1%	6,9%	3,4%

Table 1: Characterization of participant study sample

#### 4.2 Study method

The present study consisted of two main tasks. First, participants were asked to fill out a small questionnaire which inquired on participants’ ages, gender, as well as news reading preferences and behaviors. The second task required participants to visualize, during 10 minutes, news present on the SAPO Portal – Figure 1.

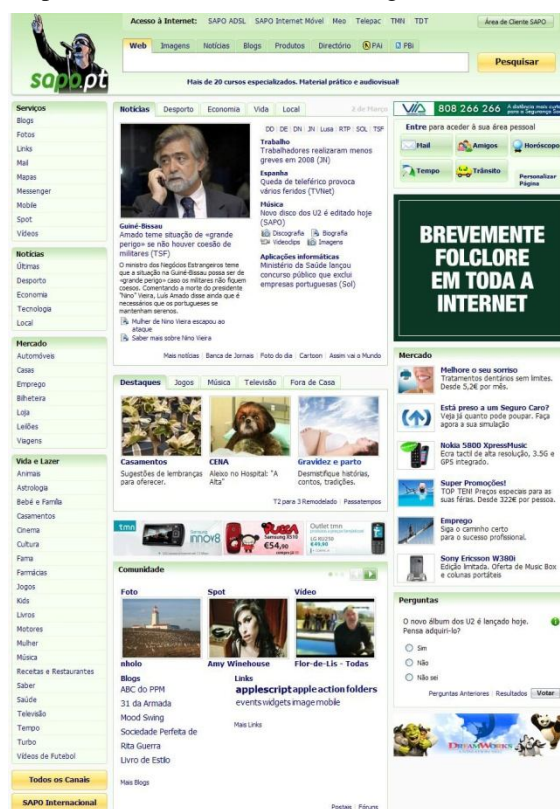


Figure 1: Representation of the SAPO home page

Specifically, participants were asked to randomly visualize any news piece present on the portal and when finished, to return to the home page and continue reading news. The supervising researcher observed each participant’s session and, if the participant began reading content that was not news, he/she was asked to return to the home page.

<sup>1</sup> Poynter Institute: <http://www.poynter.org/>



Despite the main objective of understanding participants' visualization behavior regarding news reading, several other areas of interest (AOI) on the SAPO portal were analyzed both in terms of visual behavior but also in terms of mouse interaction.

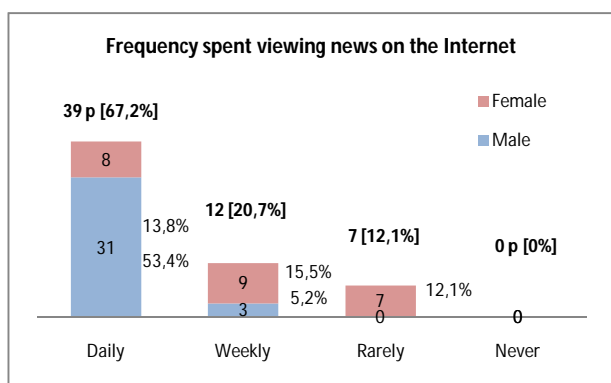
## 5. RESULTS & DISCUSSION

Data was acquired via questionnaire as well as with eye tracking software. As referred earlier in section 4.2 – Study Method – the questionnaire inquired on some personal information as well as news reading preferences.

### 5.1 Questionnaire analysis

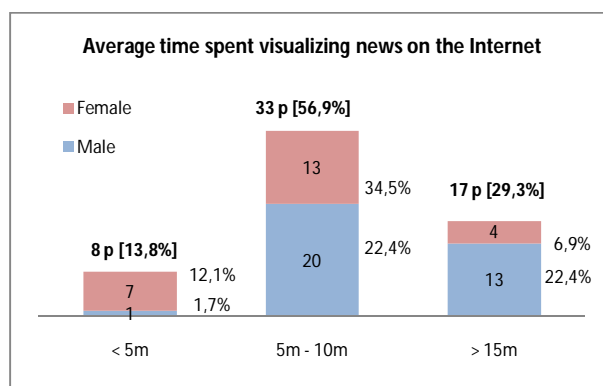
The data collected through the questionnaires helped answer three main questions: 1 – how frequently participants visualized news on the Internet; 2 – the average time participants spent reading the news on the Internet; 3 – how frequently participants accessed the SAPO portal.

Graph 1 presents the frequency participants spend viewing news on the Internet. 39 participants [p] (67.2%) indicated viewing news on a daily basis, 12 (20.7%) on a weekly basis, 7 (12.1%) indicated rarely viewing news on the internet and 0 participants answered never viewing news.



**Graph 1: Frequency of viewing news on the internet and distribution according to participant gender**

Graph 2 presents the average time, in minutes and per day, spent by the participants visualizing news on the Internet. 8 participants [p] (13.8%) indicated spending less than 5 min/day; 33 participants (56.9%) indicated spending between 5 and 15 min/day while 17 participants (29.3%) answered spending more than 15 min/day visualizing news on the Internet.



**Graph 2: Average time (in minutes) spent visualizing news on the internet and distribution according to participant gender**

The final question inquired on participants' accessing frequency of the SAPO portal. The results indicated that 5 participants (8.6%) visit the portal on a daily basis; 9 participants (15.5%) on a weekly basis; 42 participants (72.4%) rarely visit and 2 participants (3.4%) indicated never visiting the portal.

### 5.2 Visual attention data analysis

As mentioned, the second task of the study required the 58 participants to visualize news on the SAPO portal during 10 minutes. However, for the following analysis, and according to the study goal, only the time spent on the SAPO home page will be considered.

Table 2 summarizes the maximum, minimum and average time spent on the SAPO home page and percentage relative to the average duration of the participants' sessions (9m47.3s).

	Maximum	Minimum	Average
Time	5m56.5s	6sec	1m51.9s
%	57.3%	1.7%	19%

**Table 2: Maximum, Minimum and Average Time spent by participants on the SAPO home page and percentage relative to average study duration**

The following results and analysis, however, is relative to the average time spent by participants on the home page: 1min51.9sec.

#### 5.2.1 "News Area"

Figure 2 represents the "News Area". The "News Area" is divided into 4 main sections: the "Superior News Tab", "Feature Image", "Feature News piece" and "Other News".



**Figure 2: Representation of the "News Area" and the 4 main sections of the area**

Table 3 presents the average time participants spent viewing the "News Area"; approximately 1m3s, 56.3% of the average time (1m51.9s) participants spent on the SAPO home page.

	News Area	S. News Tab	Feature Image	Feature News piece	Other News
Time	1m3s	3.5sec	7.4sec	18.2sec	33.3sec
%	56.3%	3.1%	6.6%	16.3%	29.8%
Male	57.1%	2.4%	7.9%	18.1%	28.2%
Female	53.6%	3.7%	5.8%	15.5%	28.1%
<20	46.1%	3%	5.3%	17.9%	19.7%
20-29	57.6%	3%	7.3%	17.3%	29.6%
30-39	53.4%	1.6%	6.9%	12.2%	32.2%
40 +	67.6%	3.7%	11.3%	17.7%	33.4%

**Table 3: Average time participants spent visualizing the 4 sections of the "News Area" and respective percentage; average percentage spent visualizing "News Area" based on gender and age groups**

The results presented in Table 3 indicate that the section "Other News" was the most visualized area (29.8%). This value is distributed among 3-5 small news clippings that are normally present in the section. Therefore, it can be estimated that each one of these small clippings receives on average between 6-10% of visual attention; that is, a smaller value than the combination of "Feature Image" and "Feature News piece" with a total of 22.9% (16.3% + 6.6%) of participants' visual attention on the home page.

### 5.2.2 "Advertising Areas"

Four areas of advertisement were identified on the SAPO home page. Results indicate that these 4 areas had very reduced visual activity. "Advertising Area (AA) 1", the biggest of the 4 areas, received on average 2.3s of visual attention, 2% of the average time spent by participants on the SAPO home page (1m51.9s). "AA 2" received on average 1.2s (1%) and the remaining 2 areas, "AA 3 & 4" 0.1s (0.1%) of visual activity. Despite these somewhat

insignificant values, other numbers were acquired. "AA 1" was visualized by 49 participants; "AA 2" by 43 and "AA 3 & 4" by 9 and 14 participants. However, if a minimum time of visualization is defined at 1s, 31 participants visualized "AA 1", 24 "AA 2" and 2 and 1 participants "AA 3 and 4". Increasing the time variable to 5 seconds, "AA 1 & 2" was seen by 5 and 4 participants while the remaining two were not seen by any participants. Finally, despite the size and privileged location of area "AA 1" on the homepage, when compared to other areas on the homepage, the quantity of visual attention received is quite insignificant. However, it must be noted that this information should not be final due to the fact that on occasion, perceiving information present in advertisements can be done in very short periods of time. The four "Advertising Areas" found on the SAPO homepage are identified below on Figure 3.

### 5.2.3 "Side Menu"

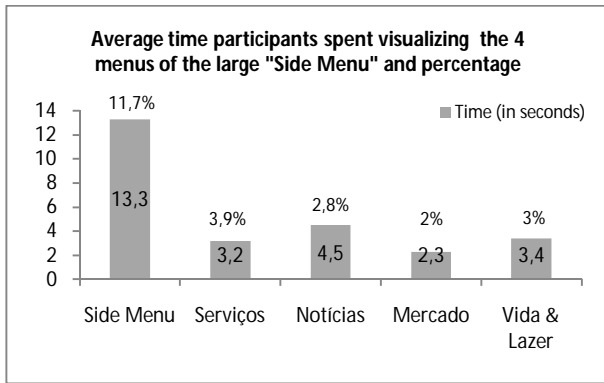
Table 4 indicates the average time participants spent visualizing the "Side Menu" and the individual menu sections: "Serviços", "Notícias", "Mercado" and "Vida & Lazer".

	Side Menu	Serviços	Notícias	Mercado	Vida & Lazer
Time	13.3s	3.2s	4.5s	2.3s	3.4s
%	11.7%	2.8%	3.9%	2%	3%

**Table 4: Average time participants spent visualizing the 4 menus of the large "Side Menu" and respective percentage**

Curiously, despite the "Menu Serviços" having less than half of the number of items present on the "Menu Vida & Lazer", the average time participants spent on either one of these menus was very similar: 3.2s (2.8%) and 3.4s (3%). A possible justification for this occurrence is related to the fact that when visualizing the home page, not all of the home page was visible. The resolution of the screen did not allow for the entire page to be displayed so a participant had to scroll down to see the entire page. Because the limit of visible area is just below the "Advertising Area 2", visible on Figure 3, some of the items on the "Menu Vida & Lazer" are "cropped" which implies that without scroll, both the "Menu Serviços" and "Menu Vida & Lazer" are approximately the same size and therefore, could receive approximately the same duration of attention.

Graph 3 represents the average time spent by participants visualizing the 4 menus – "Serviços", "Notícias", "Mercado" and "Vida e Lazer" – of the large "Side Menu" and respective percentage.



**Graph 3: Average time participants spent visualizing the 4 menus of the large “Side Menu” and respective percentage**

#### 5.2.4 Other areas of the SAPO home page

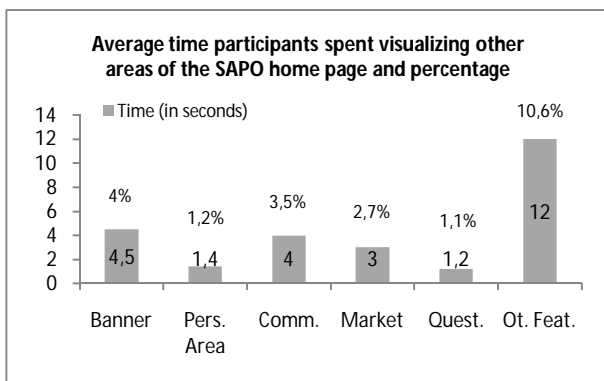
In addition to the “News Area”, “Advertising Areas” and the “Side Menu”, other areas of the SAPO home page were also analyzed. Areas such as the “Banner”, “Personal Area”, “Community”, “Market”, “Questions” and “Other Features”.

	Ban.	Per. Area	Comm.	Market	Quest.	Other Feat.
T	4.5s	1.4s	4s	3s	1.2s	12s
%	4%	1.2%	3.5%	2.7%	1.1%	10.6%

**Table 5: Average time participants spent visualizing other areas of the SAPO home page and respective percentage**

As Table 5 demonstrates, the “Other Features” area was the most viewed area (10.6%) while the “Banner” – which contains the search area – was the least viewed area with 4% of participants’ visual attention.

Graph 4 represents the average time spent by participants visualizing other areas of the SAPO home page: the “Banner”, “Personal Area”, “Community”, “Market”, “Questions”, “Other Features” and respective percentage.

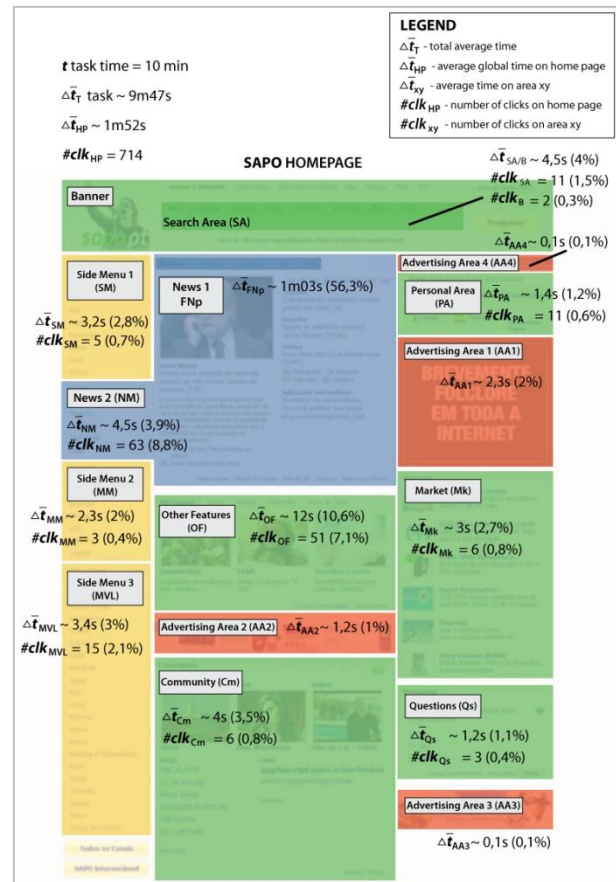


**Graph 4: Average time participants spent visualizing other areas of the SAPO home page and respective percentage**

### 5.3 Mouse interaction data analysis

Although the main purpose of the study was to understand visual behavior, participants’ mouse interaction on the areas mentioned in section 5.2 was also analyzed. Considering the 58 participants that participated in the study, a total number of 714 clicks were registered. The maximum number of clicks made by a participant was 30 whereas the minimum was 1, with an average of 12.3

clicks per participant. Furthermore, on average, 83.3% of the clicks were made inside 3 main areas: the “Side Menu”, the “News Area” and a new section, the “Search Area”. The remaining clicks were made on some of the areas mentioned in section 5.2.4.



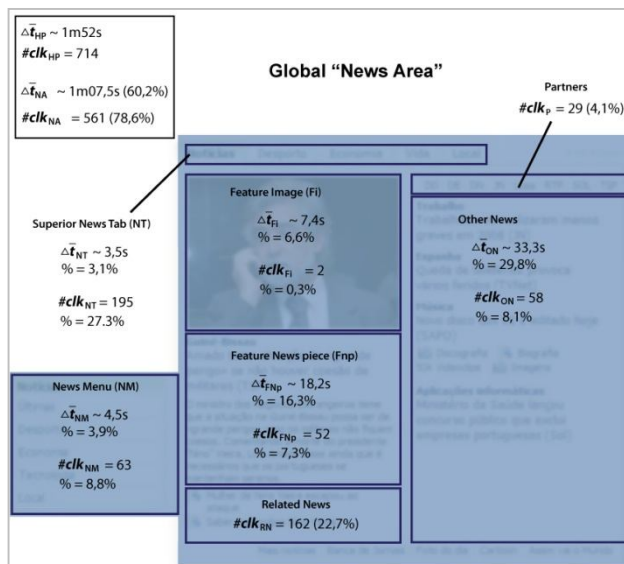
**Figure 3: Holistic SAPO home page results representing section contextualization of visual attention and average mouse clicks**

Figure 3 represents a holistic home page info-graphic overlay of the average time participants spent visualizing a pre-determined section of the homepage (results mentioned in section 5.2) as well as the number of clicks made upon those sections.

The blue area of interest is relative to the “News Area”; yellow blocks are relative to the “Side Menu”; red blocks are “Advertising Areas” and green areas are other areas of the SAPO home page.

#### 5.3.1 “News Area”

In this section, once again the “News Area” is analyzed. Contrary to what was done for section 5.2.1, two new sections are defined for the “News Area”: “Related News” and “Partners”. Furthermore, the “News menu” found on the “Side Menu” (section 5.2.3) will also be considered a part of the “News Area” for the interaction analysis. Figure 4 represents the reformulated “News Area” defined for the analysis of participants’ mouse interaction and includes: “Superior News Tab”, “Partners”, “Feature Image”, “Feature News Piece”, “Related News”, “Other News” and “News Menu”.



**Figure 4: Specific result overlay on SAPO home page News exclusive area, depicting visual attention and average mouse click for each sub-area**

Figure 4 also indicates in detail for each of the referred sections the average time participants spent visualizing the content, the relative percentage based on the total average time spent on the home page (1m51.9s), the number of clicks made in the section as well as the relative percentage based on the total number of participant clicks (714).

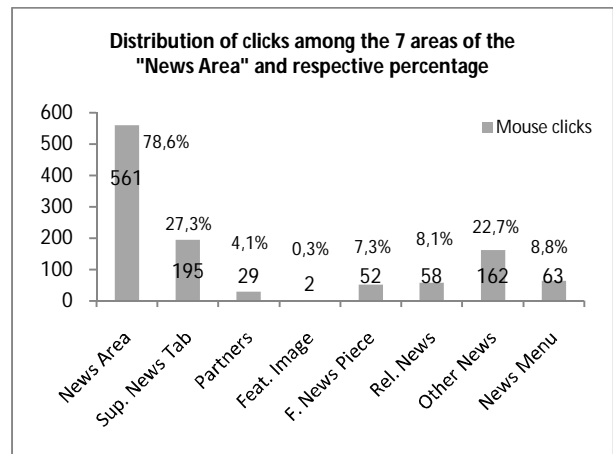
Table 6 demonstrates the distribution of clicks among the various defined areas in the "News Area". A total of 561 clicks were made in this area (78.6% of the total number of clicks – 714).

	News Area	S. News Tab	Partners	Feature Image
No. Clk	561	195	29	2
%	78.6%	27.3%	4.1%	0.3%
	Ft. News piece	Related News	Other News	News Menu
No. Clk	52	58	162	63
%	7.3%	8.1%	22.7%	8.8%

**Table 6: Distribution of number of clicks among the 7 areas of the "News Area" and respective percentage based on total number of clicks on the home page**

As the numbers on Table 6 indicate, the "Superior News Tab" received the most number of clicks, with a total of 195 whereas the "Feature Image" received only 2.

Graph 5 represents the distribution of the number of clicks among the 7 areas of the "News Area" – "Superior News Tab", "Partners", "Feature Image", "Feature News Piece", "Related News", "Other News" and "News Menu" – and respective percentage based on total number of clicks on the home page (714).



**Graph 5: Distribution of number of clicks among the 7 areas of the "News Area" and respective percentage based on total number of clicks on the home page**

The "Superior News Tab" is divided into 5 items: "News", "Sports", "Economy", "Life" and "Local". In terms of gender differences in interaction, the most notable difference resided with the "Sports" item where male participants made 35 clicks whereas female participants only made 14.

Grouping together the 4 main sections that are news related – "Feature Image", "Feature News piece", "Related News" and "Other News"; the total number of clicks registered was 274.

The "News Menu" is divided into 5 items: "Latest [news]", "Sports", "Economy", "Technology" and "Local". Of the total number of clicks made in the "News Area", only 63 were made in the "News Menu". Bearing in mind that 3 of the 5 items in this area are similar to those found in the "Superior News Tab" – "Sports", "Economy" and "Local" – and that the "Superior News Menu" received more than 3 times as many clicks (195), one can presume that participants prefer a horizontal menu positioning (and placed at a higher level) than a vertical menu placed on a page's lateral side. Towards this idea contributes the fact that despite the large difference in clicks, both "news menus" had similar average times of visualization: 3.5s (3.1%) for the "Superior Tab Menu" and 4.5s (3.9%) for the "News Menu". Additionally, of the 5 items present on the menu, the "Latest [news]" and "Technology" items were the most interacted upon with 17 and 24 clicks of the 63 total clicks. Furthermore, and although this assumption can't be conclusive, male participants interacted more upon the "Sports" and "Technology" items, a behavior some might consider a "typical" male behavior pattern.

### 5.3.2 "Side Menu"

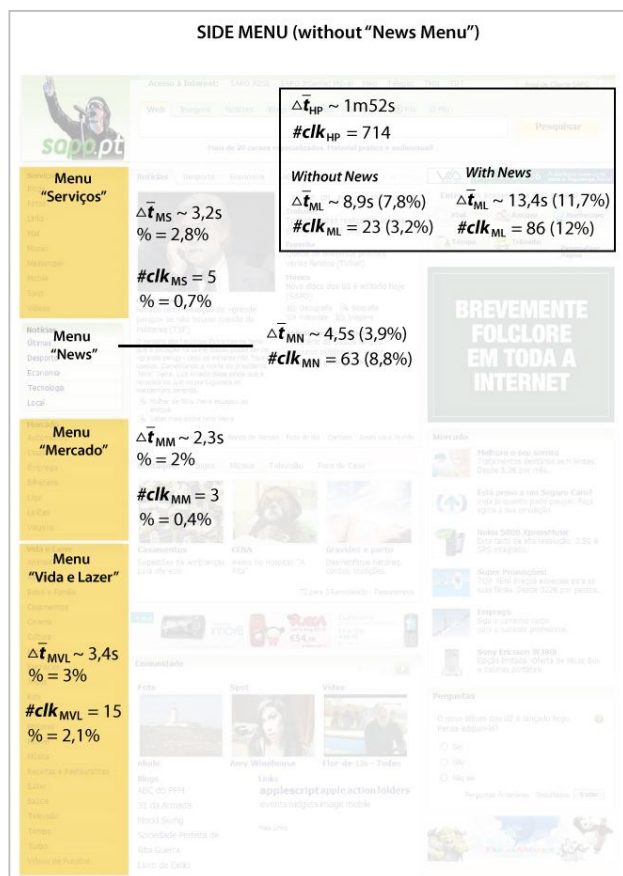
Contrary to what was done for section 5.2.3, mouse interaction analysis on the "Side Menu" will not consider interaction made upon the news-specific menu as it was analyzed previously with the "News Area" (section 5.3.1).

Figure 5 demonstrates both the average time participants spent visualizing each of the menus in the "Side Menu" (as well as the relative percentage), as well as the number



of clicks made upon the 3 menus. None of the 3 menus received a great amount of clicks as none were directly related to the objective of reading news. Nevertheless, 15 (2.1%) clicks were registered on the “Vida e Lazer Menu” – the largest of the three – 7 of which on the “Cinema” item. Furthermore, of these 7 clicks, 6 were done by participants grouped in the 20-29 age category, suggesting a greater interest in cinema from the “younger” population.

The acquired results present on Figure 6 do not indicate any significantly relevant difference between average time of participant visualization when compared to the number of clicks made on any of the menus.



**Figure 5: Specific result overlay on SAPO home page Side Menu, depicting visual attention and average mouse click for each sub-area**

### 5.3.3 Advertising and other areas of the homepage

The last of the areas that received some interaction include the “Advertising Areas” as well as other areas mentioned previously in section 5.2.4.

The total number of clicks in these areas was 119 (16.7%). 51 (7.1%) of these clicks were made on the “Other Features” section, found just below the “News Area” (in blue) in Figure 3. Other areas or items that received some relevant interaction were the “Photo of the Day” (Foto do Dia) and the “News Stand” (Banca de Jornais) items with 16 (2.2%) and 13 (1.8%) clicks, respectively. Once again, as the main objective of the study was related to news search, the lack of interaction on these areas is understandable. In terms of the “Adver-

tisement Areas”, only 3 clicks were made on any of the 4 advertising spots.

## 6. CONCLUSIONS

The present study analyzed 58 participants’ visual and interaction behavior on SAPO’s homepage. Bearing in mind the acquired results, several conclusions can be drawn regarding visual and interaction behavior.

Considering the number of clicks participants made on the horizontal tab in the “News Area”, it seems reasonable to conclude that tabs are a more efficient solution for interaction than vertical menus. Despite the average visualization time being similar, the number of clicks made on the “Superior News Tab” was more than 3 times that verified on the “News Menu”.

Additionally, tabs in general appear to be the most efficient of the interaction areas, as participants seem to have no difficulties in interacting with the tab paradigm. Furthermore, the ratio clicks/visual attention on tabs is, as the results demonstrate, superior than on any other interaction area of the SAPO homepage.

Another valuable conclusion is that in the global “News Area”, the group “Feature Image” + “Feature News piece” is more efficient than the “Other News” area where several small news pieces can be found. In fact, the results indicate that the “Feature Image” + “Feature News piece” received 4 times more visual attention than the “Other News” area.

Other acquired results reveal some interesting information regarding advertisement. The advertising areas that received the most visualizations were those placed in a favorable location. “Advertising Area 1” is favorably located next to the global “News Area”, centered on the SAPO homepage. “Advertising Area 2” is also favorably located, vertically centered on the homepage. These results verify what (Barthelson, 2002) stated regarding the importance of a prioritized layout, where content that is centered receives more attention.

While ¾ of the participants indicated that they did not, or rarely visited, the SAPO homepage, their web use literacy is rather high for Portuguese standards. The efficiency of tab use reveals that the participants have a common use of this metaphor and have adopted it comfortably. Tabs in fact provide an efficient way of organizing different contexts of information adjacently maintaining an optimization of the visual field occupation. Tab interface metaphor can also be applied with some degree of scalability and still maintain interaction design coherence.

The results acquired through the empirical study with 58 participants at the University of Aveiro complemented results acquired internally at SAPO’s usability laboratory, and had a direct consequence on the reconception and redesign of the SAPO homepage. SAPO’s current homepage uses tabs as a dominant scheme of organization and interaction. Additionally, the side menu was removed, freeing up vertical space used for information display. The areas previously indexed in the side menu are now presented as boxes embedded in the page. Finally, areas

exclusive to advertising were not changed and remain in the same place.

Figure 6 represents SAPO's current homepage. At the top of the homepage it is clear that a horizontal navigation paradigm is dominant, as the 4 red outlined boxes indicate. The two highlighted blue boxes at the bottom represent the reorganization of the side menus into "boxes" with items presented vertically. Furthermore, as mentioned, there are no traces of the previous side menu.



Figure 6: Current SAPO homepage layout

## 7. ACKNOWLEDGEMENTS

We would like to thank SAPO/UA R&D Labs at University of Aveiro for sponsoring this research and Benjamin Júnior, Eduardo Pinto, Ivo Gomes and Rui Gomes from the Usability Lab at SAPO headquarters in Lisbon for letting us use their eye tracker in this study.

Special thanks to all the students that participated in the study and sponsored it with their time and dedication.

To Mário Rodrigues, colleague and friend at the Department of Communication and Art of the University of Aveiro, thank-you for your precious support setting up the eye tracking equipment and all the software and hardware configuration that was needed.

## 8. REFERENCES

- Barthelson, M. (2002). *Behaviour in online news reading*: Lunds Universitet.
- Duchowski, A. (2007). *Eye Tracking Methodology: Theory and Practice* (Second Edition ed.): Springer.

Garcia, M. (1991). *Eyes on the News*: The Poynter Institute for Media Studies.

Holmqvist, K., Holsanova, J., Barthelson, M., & Lundqvist, D. (2003). Reading or Scanning? A Study of Newspaper and Net Paper Reading. In R. Radach, J. Hyona & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*: Elsevier.

Hsieh-Yee, I. (2001). Research on Web search behavior. [doi: DOI: 10.1016/S0740-8188(01)00069-X]. *Library & Information Science Research*, 23(2), 167-185.

Jacob, R. J. K., & Karn, K. S. (2003). Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises. *The Mind's eye: Cognitive The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, 573-603.

Leggett, D. (January19, 2010). A Brief History of Eye Tracking. *UX Booth* Retrieved April 28, 2010, from <http://www.uxbooth.com/blog/a-brief-history-of-eye-tracking/>

Lewenstein, M., Edwards, G., Tatar, D., & DeVigal, A. (2000). Stanford Poynter Project. from <http://www.poynterextra.org/et/i.htm>

Mosconi, M., Porta, M., & Ravarelli, A. (2008). *On-line newspapers and multimedia content: an eye tracking study*. Paper presented at the Proceedings of the 26th annual ACM international conference on Design of communication.

Nielsen, J., & Pernice, K. (2009). *Eyetracking Web Usability*. Berkeley, CA: New Riders Press.

Outing, S., & Ruel, L. (2004). Eyetrack III. from <http://www.poynterextra.org/eyetrack2004/main.htm>

Pan, B., Hembrooke, H. A., Gay, G. K., Granka, L. A., Feusner, M. K., & Newman, J. K. (2004). *The determinants of web page viewing behavior: an eye-tracking study*. Paper presented at the Proceedings of the 2004 symposium on Eye tracking research & applications.

Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. [doi: DOI: 10.1016/j.visres.2007.05.007]. *Vision Research*, 47(21), 2714-2726.



# Interfaces para Hipervídeo em 360°

Luís A. R. Neng  
LaSIGE, Faculdade de Ciências  
Universidade de Lisboa  
+351.21.750 00 87  
luisarn@gmail.com

Teresa Chambel  
LaSIGE, Faculdade de Ciências  
Universidade de Lisboa  
+351.21.750 00 87  
tc@di.fc.ul.pt

## Sumário

*Neste artigo descrevemos uma abordagem para o design e desenvolvimento de uma interface imersiva e interactiva para a visualização e navegação de hipervídeos em 360° através da internet. Este tipo de hipervídeos permite aos utilizadores movimentarem-se em torno de um eixo para visualizarem os conteúdos dos vídeos em diferentes ângulos e acederem-lhes de forma eficiente através de hiperligações. Desafios para a apresentação deste tipo de hipervídeos incluem: proporcionar aos utilizadores uma interface adequada que seja capaz de explorar conteúdos em 360° num ecrã normal, onde o vídeo deve mudar de perspectiva para que os utilizadores sintam que estão a olhar ao redor, e formas de navegação adequadas para compreenderem facilmente a estrutura do hipervídeo, mesmo quando as hiperligações estejam fora do alcance do campo de visão.*

## Palavras-chave

*Hipervídeo, Hipermédia, Hiperligação, 360°, Navegação, Streaming de vídeo, Vídeos interactivos, Hotspots, Interfaces imersivas.*

## 1. INTRODUÇÃO

Os dispositivos para a captura de vídeo em 360°, bem como as formas de os disponibilizar na Web, são cada vez mais comuns e acessíveis ao público em geral. Neste contexto, é pertinente explorar formas e técnicas de navegação para visualizar e interagir com hipervídeos em 360° na Web. Tradicionalmente, para visualizar o conteúdo de um vídeo, o utilizador fica limitado à região para onde a câmara estava apontada durante a sua captura, o que significa que o vídeo resultante terá limites laterais. Com a gravação de vídeo em 360°, já não há estes limites: abrindo novas direcções a explorar. Um *player* de hipervídeo em 360° vai permitir aos utilizadores movimentarem-se à volta para visualizar o resto do conteúdo e aceder de forma fácil às informações fornecidas pelas hiperligações. O vídeo é um tipo de informação muito rico que apresenta uma enorme quantidade de informação que muda ao longo do tempo. Um vídeo em 360° apresenta ainda mais informações ao mesmo tempo e acrescenta desafios, pois nem tudo está dentro do nosso campo de visão. No entanto, proporciona ao utilizador uma nova experiência de visualização potencialmente imersiva.

Exploramos técnicas de navegação para ajudar os utilizadores a compreenderem e navegarem facilmente um espaço de hipervídeo a 360° e proporcionar uma experiência de visualização a outro nível, através dum espaço hipermédia imersivo. As hiperligações levam o utilizador para outros conteúdos hipermédia relacionados, tais como textos, imagens e vídeos ou outras páginas na web. Depois de terminar a reprodução ou visualização dos conteúdos relacionados, o utilizador poderá retornar à

posição anterior no vídeo. Através da utilização de técnicas de sumarização, podemos ainda fornecer aos utilizadores um sumário de todo o conteúdo do vídeo para que possam visualizá-lo e compreendê-lo numa forma mais eficiente e flexível, sem necessitar de visualizar o vídeo todo em sequência.

Nas próximas secções, descrevemos alguns formatos de vídeo em 360° e diferenças entre hipervídeo e hipervídeo em 360°. Em seguida, apresentamos um *player* para visualização de hipervídeos em 360°, focando essencialmente o design de elementos de navegação, e depois faremos uma revisão de trabalhos relacionados. Por fim, apresentamos algumas conclusões obtidas e perspectivas para trabalho futuro.

## 2. VÍDEOS EM 360°

O vídeo tem provado ser uma das formas mais eficientes de comunicação, permitindo a apresentação de um leque enorme e variado de informação num curto período de tempo. Os vídeos em 360° podem fornecer ainda mais informação, podendo ser mapeados sobre projecções cilíndricas ou esféricas. A Figura 1a apresenta o exemplo de uma *frame* de um vídeo capturado através da câmara Sony Bloggie PM5K [Sonya]. Na Figura 1b, o mesmo vídeo é projectado de forma panorâmica, mostrando todo o ângulo de 360° [Sonyb]. O exemplo da Figura 2 é um passeio de mota capturado e projectado com os mesmos mecanismos da Figura 1, permitindo visualizar ao mesmo tempo o próprio condutor e a estrada que se encontra à sua frente [KTM]. A projecção cilíndrica foi inventada em 1796 pelo pintor Robert Barker de Edimburgo que obteve a sua patente [Yelick80].

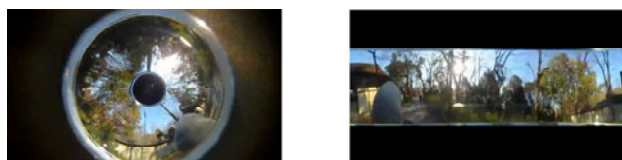


Figura 1: a) Vídeo cilíndrico antes de ser projectado;  
b) Vídeo cilíndrico depois de ser projectado



Figura 2: Passeio de mota num vídeo a 360°

### 3. HIPERVÍDEO E HIPERVÍDEO EM 360°

A utilização de vídeo na Web tem consistido essencialmente na sua inclusão nas páginas, onde são visualizados de forma linear, e com interações em geral limitadas às acções de *play* e *pause*, *fast forward* e *reverse*. Nos últimos anos, os avanços mais promissores no sentido do vídeo interactivo parecem ser através de hipervídeo, proporcionando uma verdadeira integração do vídeo em espaços hipermedia, onde o conteúdo pode ser estruturado e navegado através de hiperligações definidas no espaço e no tempo e de mecanismos de navegação interactivos flexíveis [Chambel02]. Ao estender o conceito de hipervídeo para 360°, surgem novos desafios, principalmente porque grande parte do conteúdo está fora do campo de visão. O *player* de hipervídeo a 360° tem que fornecer aos utilizadores mecanismos apropriados para facilitar a percepção da estrutura do hipervídeo, para navegar de forma eficiente no espaço hipervídeo a 360° e idealmente proporcionar uma experiência imersiva.

### 4. VISUALIZAÇÃO DE HIPERVÍDEO EM 360°

Para poder navegar num espaço hipervídeo a 360°, necessitamos de novos mecanismos de navegação. Esta secção apresenta os principais mecanismos concebidos para visualização deste tipo de hipervídeo e soluções para os principais desafios em hipermedia: desorientação e sobrecarga cognitiva, agora no contexto de 360°. Focaremos, essencialmente, os mecanismos de navegação que ajudam o utilizador a orientar-se no espaço de 360°.

#### 4.1 Interface de Navegação no Vídeo

Desenvolvemos uma interface que funciona por arrastamento, Figura 3, para a navegação no vídeo em 360°. Esta interface permite que o utilizador movimente o vídeo para visualizar o conteúdo em diferentes ângulos. O utilizador só precisa de arrastar o cursor para a esquerda ou para a direita para movimentar o campo de visão. Pode no entanto movimentar-se apenas para um dos lados para dar a volta sem qualquer tipo de limitação.

A percepção da localização e do ângulo de visualização actual tornou-se um problema devido à falta de limites laterais. Durante os nossos testes, muitos utilizadores sentiram-se perdidos no espaço de 360°, sem saber que ângulos é que estavam a visualizar. Este foi um dos aspectos abordados nos mecanismos de navegação criados, e apresentados em seguida.

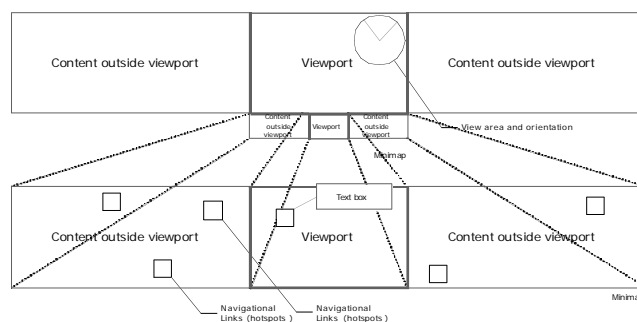


Figura 3. Maqueta do Player de Hipervídeo em 360°

#### 4.2 Navegação no Hipervídeo a 360°

Desenvolvemos dois tipos de mecanismos de navegação para ajudar os utilizadores a se **localizarem** dentro do espaço de vídeo em 360°.

- **View area:** é um círculo parecido com um gráfico circular onde existe uma única “fatia” que corresponde ao ângulo de visualização em que o utilizador está posicionado. Esta roda para a esquerda ou para a direita, consoante o lado para o qual o utilizador está a arrastar, permitindo a sua orientação.
- **Mini map:** é uma projecção cilíndrica do vídeo redimensionado para um tamanho que possa ser visualizado por completo na interface. O cursor do *mini map* – um rectângulo sobre o *mini map* (Figura 4) – indica que parte do vídeo é que o utilizador está a visualizar naquele momento.

Em hipervídeo, a **percepção de hiperligações** é mais desafiante do que em hipermedia tradicional porque as hiperligações podem ter duração, podem coexistir no tempo e no espaço e o vídeo muda ao longo do tempo. Assim, são precisos mecanismos especiais, para torná-las perceptíveis aos utilizadores, [Chambel02]. Em hipervídeo em 360°, grande parte do conteúdo é invisível ao utilizador por não estar no campo de visão, logo será necessário estudar novas abordagens e mecanismos para indicar a existência de hiperligações.

Criámos os **hotspots availability** e **location indicators** (Figura 4) para permitir aos utilizadores saberem a existência e a localização de cada uma das hiperligações. O posicionamento dos indicadores de **hotspots availability** no eixo da ordenada, nas margens laterais do vídeo, serve para indicar em que posição vertical está cada uma das hiperligações. O tamanho do indicador serve para indicar a distância do **hotspot** em relação ao ângulo de visualização. Quanto mais perto fica o **hotspot**, maior é o indicador. Os indicadores são semi-transparentes e estão posicionados nas margens laterais para minimizar o impacto que têm sobre o conteúdo do vídeo.

O **Mini Map** também fornece informações acerca da existência e localização de hotspots (Figura 4), que deverão conter alguma informação do conteúdo de destino, para que o utilizador possa ter alguma expectativa acerca do que vai visualizar depois de seguir a hiperligação. Uma caixa de texto com aspecto de balão de banda desenhada permite acomodar várias informações relevantes. Por exemplo, na Figura 4, colocámos a informação acerca do nome, descrição e preço do capacete.

Quando os utilizadores *seleccionam o hotspot*, poderão ser redireccionados para uma página com info adicional ou a selecção pode ser memorizada pelo sistema e o seu conteúdo ser mostrado apenas quando o utilizador desejar, dependendo do tipo de aplicação. Por exemplo, se a finalidade do vídeo for o apoio à aprendizagem (*e-learning*), pode fazer mais sentido abrir logo o conteúdo da hiperligação, pois os utilizadores estão habituados a ver aquele tipo de informação passo a passo. Se o vídeo for de entretenimento, os utilizadores provavelmente não gostam de ser interrompidos pela abertura do novo conteúdo, podendo optar pela memorização da hiperligação, e pelo seu acesso posterior, quando quiserem.

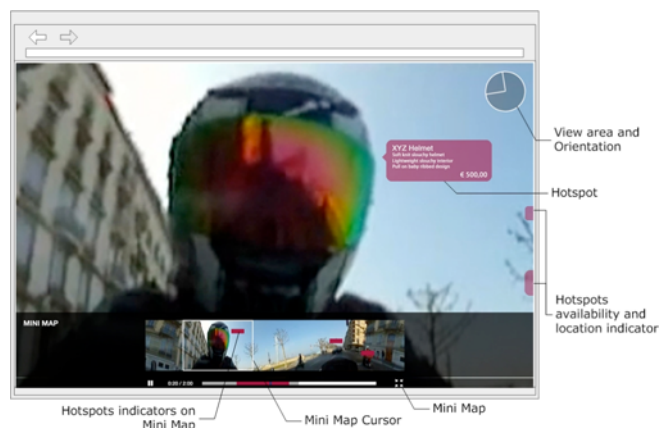


Figura 4. Player de Hipervídeo em 360° em modo Mini Map

O Player de Hipervídeo em 360° fornece outro modo de visualização (Figura 5). Para além do título e da descrição do vídeo, o modo **Video Image Map** fornece uma visualização global do conteúdo do vídeo. É o mesmo tipo de **sumarização** descrita em [Chambel02, Shipman03] mas com algumas funcionalidades extra e adaptações ao espaço 360°. As pré-visualizações (*thumbnails*) referem-se às cenas do vídeo e são representadas através duma projecção cilíndrica, para que todo o conteúdo ao longo do tempo possa ser visualizado. Permite também, de forma sincronizada, saber a cena actual e oferece ao utilizador a possibilidade de navegar para outras cenas. Toda a área de pré-visualização é sensível ao clique e determina as coordenadas da pré-visualização que o utilizador seleccionou. Por exemplo, se o utilizador seleccionar a parte direita do terceiro *thumbnail*, será redireccionado para o ângulo do vídeo seleccionado e no tempo 00:00:20, a que corresponde aquele *thumbnail*. Uma versão mais condensada disponibiliza apenas a pré-visualização da parte central de cada uma das cenas. Permite a apresentação simultânea de um maior número de cenas, mas limita a visualização e a flexibilidade para navegar para o ângulo desejado de forma mais directa.

Algumas funcionalidades também foram adicionadas à linha de tempo (*timeline*), ou **Barra de Progresso**, tradicionais (Figura 6). Para além dos tradicionais botões de *Play*, *Pause* e Tempo de Vídeo, estendemos a barra para adaptar a algumas características de uma página Web. Como é um Player desenvolvido para funcionar na internet, precisamos de ter em conta que é preciso tempo para carregar o vídeo. A barra de **bytes loaded** indica ao utilizador o progresso do carregamento do vídeo e não

zador o progresso do carregamento do vídeo e não permite que o utilizador acesse às informações que ainda não foram carregadas.

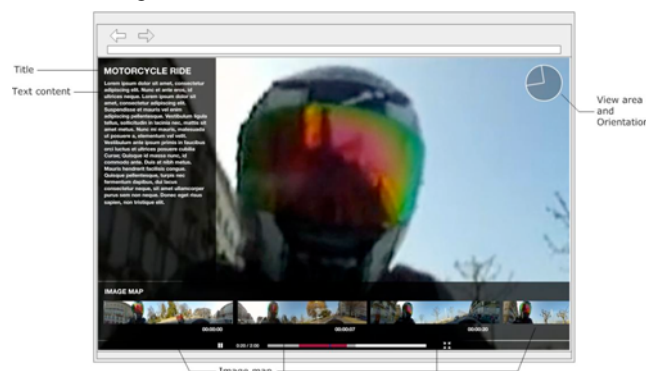


Figura 5. Player de Hipervídeo a 360° em modo Image Map

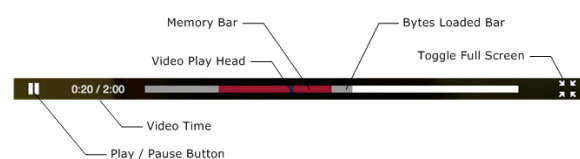


Figura 6. Barra de progresso e controlo do Player Hipervídeo em 360°

O hiperespaço é navegado em contextos espaço-temporais que a história recorda [Chambel02]. A barra de memória, **Memory Bar**, fornece informação ao utilizador acerca das partes do vídeo que já foram visualizadas. O botão **Toggle Full Screen** alterna o modo de visualização do vídeo entre *full* e *standard screen*. O modo *full screen* leva o utilizador para fora das limitações do browser e maximiza o conteúdo do vídeo para o tamanho do ecrã. É mais um passo para um modo de visualização imersiva, por ex. numa projecção 360° dentro duma Cave, como estamos a considerar explorar em trabalho futuro.

## 5. TRABALHO RELACIONADO

O conceito de hipervídeo já existe desde a época inicial do hipertexto, quando Ted Nelson estendeu o seu modelo hipermédia para incluir “*branching movies*” ou “*hyperfilms*”, mas não foi adoptado globalmente como se esperava, devido a limitações essencialmente tecnológicas [Chambel02, Tiellet10]. Quando o vídeo começou a ser mais fácil de processar, armazenar e transmitir através de redes de banda-larga, testemunhamos um crescimento rápido da utilização de vídeos através da internet, tornando mais relevante adoptar o hipervídeo e criar ferramentas que suportem a sua edição e acesso.

O vídeo em hipermédia foi discutido em finais dos anos oitenta e no início dos anos noventa [Chambel02, Tiellet10]. Desde meados dos anos noventa, alguns trabalhos como a Kon-Tiki e HyperCafe incluíram hiperligações em vídeo, e outros, como o Hyper-Hitchcock [Shipman03], exploravam o vídeo “*detail-on-demand*”, focando na sumarização e criação de conceitos de hipertexto espacial. Hyperssoap abordou o hipervídeo para o contexto de entretenimento e de televisão interactiva, por exemplo, atrasando a abertura do conteúdo da hiperligação até ao final dos programas. A linguagem W3C SMIL, basea-

da no modelo AHM, deu um passo importante na integração e sincronização multimédia na Web, mas o suporte a vídeo era ainda limitado. Nos finais dos anos noventa, algumas iniciativas comerciais tiveram sucesso moderado. No contexto do suporte à aprendizagem, as *Web lectures*, trabalhos mais recentes na área de análise colaborativa de vídeo suportada por computador, e o nosso próprio trabalho anterior [Chambel02,Tiellet10] contribuíram para o desenvolvimento desta área. No entanto, a visualização e a interacção com hipervídeo em 360° é uma área de investigação nova. Existe um número elevado de aplicações e ferramentas para visualizar fotografias panorâmicas, mas existe pouco trabalho sobre vídeo em 360° e ainda menos sobre hipervídeo em 360°.

Google Street Maps [GoogleSM] pode ser considerada uma das aplicações mais conhecidas para a visualização de fotos em 360°. O utilizador pode movimentar-se para visualizar o conteúdo em diferentes ângulos e “mergulhar” na cena para seguir as ruas disponíveis. Utiliza a projecção esférica e proporciona aos utilizadores uma sensação imersiva mas não tem suporte a vídeos. A página do YouTube pode ser considerada como uma das aplicações mais populares de vídeo, e hipervídeo quando começaram a suportar anotações nos vídeos [YouTubeAnn], funcionando como hiperligações, não em 360°. Immersive Media [ImMedia] é um fornecedor de vídeos interactivos em 360°. A câmara de alta-resolução desenvolvida por eles está a ser utilizada nas áreas de entretenimento, publicidade, mapas, etc. Desenvolveram também um *Player* de Vídeo a 360°, mas neste momento ainda não tem suporte a hiperligações. A última versão da câmara Bloggie Handy Cam da Sony é uma das câmaras 360° disponíveis para o consumidor em geral, e, através da aplicação Sony Picture Motion Browser, podemos converter o vídeo numa projecção cilíndrica, mas também não tem qualquer suporte a hiperligações. Quake III Arena 360 Panoramic Virtual Reality Installation with PanoramaScreen [Quake360] é uma projecção cilíndrica sobre um espaço cilíndrico, desenvolvido pelo Institute for Visual Media [ZKM] em colaboração com o iCinema Research Centre [iCinema], Sydney, desde 2005. Partilha várias funcionalidades e desafios do hipervídeo a 360°, mas o foco principal é a realidade virtual e a interacção com jogos.

## 6. CONCLUSÕES E TRABALHOS DO FUTURO

Neste artigo, apresentamos uma abordagem para a visualização e interacção de vídeos em 360°. A navegação num espaço de vídeo em 360° apresenta uma nova experiência para grande parte das pessoas e não existem ainda intuições consistentes sobre o comportamento deste tipo de navegação. Os utilizadores, muito provavelmente, vão sentir o problema que inicialmente houve com o hipertexto, em que o utilizador se sentia perdido no hiperespaço. Por isso, o *Player* de Hipervídeo a 360° tem que ser o mais claro e eficaz possível para que os utilizadores possam interagir facilmente.

O trabalho futuro consiste na realização de mais testes de usabilidade e desenvolvimento de diferentes versões do *Player* de Hipervídeo em 360°, com mecanismos de nave-

navegação revistos e estendidos, com base nos resultados das avaliações. O *Player* de Hipervídeo em 360° não deverá ser apenas uma aplicação para Web, deverá poder integrar com quiosques multimédia ou outras instalações imersivas. Provavelmente serão necessárias novas funcionalidades e tipos de navegação para adaptar a diferentes contextos. O exemplo do *Player* de Hipervídeo em 360° apresentado neste artigo utiliza um Web browser e um rato como meio de apresentação e interacção. Com o crescimento das tecnologias de vídeo 3D, multi-toque e *eye-tracking*, podem surgir novas formas de visualização e de interacção com o espaço 360°. Estas novas formas trazem novos desafios mas também um potencial acrescido de novas experiências a explorar.

## 7. AGRADECIMENTOS

Queremos agradecer aos autores dos vídeos em 360° utilizados como exemplos. O trabalho apresentado neste artigo foi parcialmente financiado pelo LaSIGE através do Programa de Financiamento Plurianual da FCT.

## 8. REFERÊNCIAS

- [Chambel02] Chambel, T., and Guimarães, N. Context Perception in Video-Based Hypermedia Spaces. In Proc. of ACM Hypertext'2002, MD, USA.
- [GoogleStrMaps] Google Street Maps.  
[www.google.com/intl/en\\_us/help/maps/streetview/](http://www.google.com/intl/en_us/help/maps/streetview/)
- [iCinema] iCinema Research Centre, Sydney.  
[www.icinema.unsw.edu.au/](http://www.icinema.unsw.edu.au/)
- [ImMedia] Immersive Media: [www.immersivemedia.com](http://www.immersivemedia.com)
- [KTM] Test ride Sony Bloggie 360° KTM SMT  
[www.youtube.com/watch?v=OUbMvmagACc](http://www.youtube.com/watch?v=OUbMvmagACc)
- [Nelson74] Nelson, T. Branching presentational systems-Hypermedia, Dream Machines, 1974, 44-45.
- [Quake360] Quake III Arena in a 360 Panoramic VR Installation with PanoramaScreen, game demo.  
[www.youtube.com/watch?v=Cqh\\_F4iN8pg](http://www.youtube.com/watch?v=Cqh_F4iN8pg)
- [Shipman03] Shipman, F., Girgensohn, A., Wilcox, L. Generation of Interactive Multi-Level Video Summaries. In Proc. of ACM MM'2003. pp.392-401.
- [Sonya] Sony Bloggie “360 Panorama Lens” sample  
[www.youtube.com/watch?v=gG2gzWHm09o](http://www.youtube.com/watch?v=gG2gzWHm09o)
- [Sonyb] Sony Bloggie “360 Panorama Lens” Panorama expanded sample [www.youtube.com/watch?v=Nsp-45rmxb4](http://www.youtube.com/watch?v=Nsp-45rmxb4)
- [Tiellet10] Tiellet, C.A.B., Pereira, A.G., Reategui, E.B., Lima, J.V., Chambel, T. Design and Evaluation of a Hypervideo Environment to Support Veterinary Surgery Learning. In Proc. of ACM Hypertext'2010, Toronto, Canada.
- [Yelick80] Yelick, S. Anamorphic Image Processing, MIT, May 1980. <http://dspace.mit.edu/bitstream/handle/1721.1/42957/08902437.pdf?sequence=1>
- [YouTubeAnn] YouTube annotations:  
[www.youtube.com/t/annotations\\_about](http://www.youtube.com/t/annotations_about)
- [ZKM] ZKM | Institute for Visual Media, Germany.  
<http://on1.zkm.de/zkm/e/institute/bildmedien/>

# The last words of Domenico: um instrumento de performance para a reconstrução sonora de discursos

Vitor Lago Silva  
Laboratório de Criação Digital  
Matosinhos  
victorls@yahoo.com

Pedro Branco  
engageLab / Dep. de Sistemas de Informação  
Universidade do Minho, Guimarães  
pbranco@dsi.uminho.pt

## Sumário

*Este artigo apresenta, a criação e o estudo de um instrumento para performance - Fato Sensível Wireless - constituído por um fato e uma luva equipados com sensores, que detectam os movimentos de um performer. A construção deste instrumento conduziu à criação e à apresentação pública de um projecto - The last words of Domenico, que, se baseia na relação entre o movimento e o discurso. Através do movimento de um performer, geram-se conteúdos verbais sonoros modelados em tempo real. É ainda apresentado o resultado de uma entrevista para avaliação do instrumento de performance por um coreógrafo.*

## Palavras-chave

*Arte digital, performance, instrumento, sensores.*

## 1. INTRODUÇÃO

O projecto *the last words of Domenico* consiste numa performance onde se reconstrói sonoramente um discurso a partir dos movimentos livres do *performer*, com a aplicação de várias texturas ao resultado sonoro. O intuito principal é o de traduzir o gesto para um discurso sonoro.

O instrumento para performance consiste num fato e luva desenvolvidos com sensores e componentes electrónicos que permitem fazer a leitura dos movimentos das pernas e dos braços de um *performer*, bem como os movimentos do seu corpo.

O discurso é gerado a partir de um discurso original pré-existente, que é desconstruído e, depois, reconstruído através do gesto do *performer*.

É aqui discutido o sistema desenvolvido assim como relatado a experiência de um dançarino profissional na sua avaliação.

## 2. MOTIVAÇÃO

As influências e inspirações para este trabalho provêm de vários trabalhos artísticos nas áreas da performance, instalação e vídeo. O trabalho aqui proposto procura inserir-se em alguns aspectos da corrente de pensamento pós-humanista. As origens deste pensamento remontam a Nietzsche e ao conceito de super-homem, que visa a superação da espécie humana por si própria. O ser pós-humano será aquele que possui princípios de auto-formação e um carácter transcendente, porque será potencialmente imortal - seja enquanto cyborg ou trans-humano [Rüdiger07].

Actualmente, as tecnologias de realidade virtual, a engenharia genética, a medicina restauradora, as operações de mudança de sexo, as próteses de todos os tipos insinuam, que já está em curso um processo profundo de alterações e que é, ao mesmo tempo, bastante perturbador no que se refere à condição humana.

O projecto aqui proposto enquadra-se, não nas ideias radicais trans-humanistas de perpetuação da vida, mas sim na ideia da extensão física do corpo - *cyborg*. Uma extensão, não para tornar o corpo mais forte ou mais resistente a doenças, mas para dotá-lo de mais um meio de expressão - traduzir o gesto para um discurso sonoro aleatório. Não é pretendido desenvolver um meio de comunicação verbal objectivo, que traduza o pensamento humano, mas sim uma forma de construção de discurso verbal poético, subjectivo e musical através do gesto.

O trabalho de Sterlac tornou-se uma das fortes influências para este projecto, pela sua singularidade e pela forma como ele utiliza a tecnologia como extensão do seu próprio corpo. Sterlac afirma que o corpo humano se encontra obsoleto e, que, por isso, necessita de upgrades tecnológicos [Elliot]. A influência do seu trabalho, não se revela tanto ao nível da extensão tecnológica do corpo, mas sobretudo na ideia de que o corpo pode sustentar instrumentos tecnológicos que permitem novas formas de expressão.

O trabalho Nelken, de Pina Bausch, [Bausch82], teve uma influência no que se refere à questão da gestualidade inerente a este projecto. A um dado momento do espectáculo, um *performer* traduz a música *The man I love* de Gershwin em linguagem gestual.



O projecto, *Perfect Human* [Satomi08] apresenta um sistema semelhante ao usado neste projecto onde um fato com sensores faz a leitura de múltiplas articulações de um *performer*. Este foi um exemplo encontrado, de construção de um sistema de captura de movimentos economicamente acessível.

### 3. THE LAST WORDS OF DOMENICO

A performance *The last words of Domenico* baseia-se na reconstrução através do movimento de um discurso de um louco que reivindica, numa performance de rua, os seus direitos - o direito à manifestação de opinião e defesa de igualdade de direitos das pessoas consideradas loucas pela sociedade. O louco é a personagem Domenico do filme *Nostalghia* [Tarkovsky83]. A selecção deste discurso deve-se ao facto do mesmo possuir um carácter de intervenção política, social e filosófica.

O discurso originalmente em italiano foi traduzido para inglês e, posteriormente, dividido em amostras mais pequenas, correspondentes a secções do início, meio e fim da frase, de forma a poder recriar aleatoriamente um novo discurso, tendo como pressuposto que o gerar desse novo discurso possa revelar significados ocultos do discurso original de Domenico.

Desta forma, as palavras de Domenico voltariam a ser ouvidas através de novos sentidos. Dar nova voz a Domenico tornou-se, também, um dos objectivos deste trabalho.

### 4. IMPLEMENTAÇÃO

O instrumento para performance consiste num fato onde se instalaram componentes electrónicas que permitem fazer a leitura dos movimentos das pernas e dos braços do *performer*, bem como os movimentos do seu corpo. Foi ainda construída uma luva, para a leitura do movimento dos três dedos da mão esquerda. O instrumento envia via rádio os dados dos sensores para um computador de forma a dar liberdade de acção ao *performer*.



Figura 1: Fato sensível wireless

O fato é construído a partir de um fato protector para condutores de motociclos (fig. 1). Além da opção estética

do autor, este foi escolhido por razões de protecção, a zona das costas possui uma rigidez que permite manter seguras as componentes electrónicas quando o *performer* se movimenta.

A luva foi adaptada a partir de uma luva de golfe de forma a serem incorporados sensores de flexão. O uso de um capacete a cobrir a cara (fig. 2) permite omitir a identificação do *performer* e ao mesmo tempo despersonalizar-lo ao esconder as suas expressões faciais. Por outro lado, o capacete já tinha sido utilizado num trabalho prévio de 1998, *Obsessões*, no qual surgia como um *ready-made*, tornando-se uma referência pessoal do autor.



Figura 2: Performer usando o fato sensível wireless

#### 4.1 Hardware

O componente do hardware do Fato Sensível Wireless é constituído por vários blocos: sensores de flexão e aceleração, ligação dos sensores ao microprocessador Arduino, sistema de comunicação sem fios (Xbee) e alimentação de energia.

O total dos sensores instalados no Fato foram onze – quatro *bends*, dois para os braços e pernas, que fazem a leitura das suas principais flexões; três *bends* na luva, que fazem a leitura da flexão de três dedos; um acelerómetro que faz a leitura do movimento de aceleração do corpo nos três eixos X, Y, Z; e um botão *on/off*.

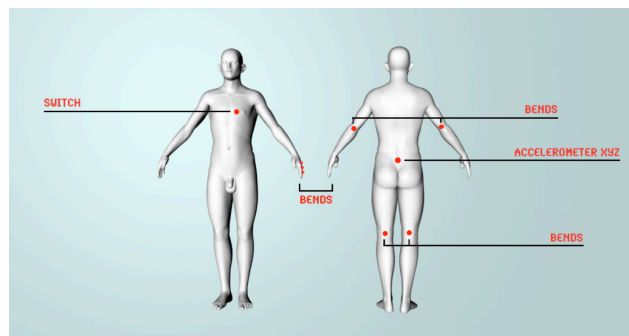
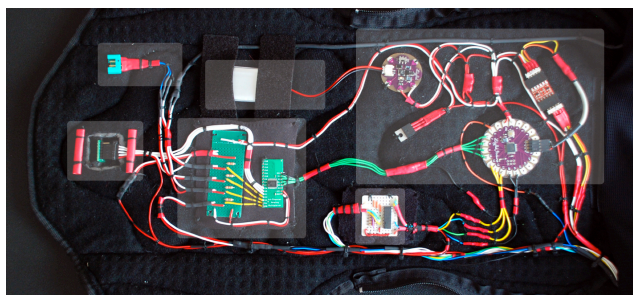


Figura 4: Posicionamento dos sensores

A figura 4 ilustra o posicionamento dos sensores em relação ao corpo do performer e a figura 5 ilustra parte do sistema montado nas costas do fato.





**Figura 5: Parte do sistema electrónico montado nas costas do fato.**

Neste sistema de leitura de movimentos não existe a distinção entre diferentes movimentos de um braço ou perna, bem como a velocidade dos mesmos. A intenção era apenas de detectar movimentos dos membros e corpo e não traduzir movimentos específicos.

#### 4.2 Software

A par da criação, construção e teste do fato sensível wireless foi desenvolvida a componente de Software que suporta a reconstrução sonora do discurso a partir dos movimentos do performer. Foram usadas as ferramentas GhostReader, Max-msp e Live.

Após a tradução do discurso de Domenico no filme *Nostalgia* de Andrei Tarkovsky de italiano para inglês utilizou-se o GhostReader [ConvenienceWare] um software *text-to-speech*, para transferir o discurso escrito de Domenico para o discurso sonoro. A voz utilizada nessa tradução é a voz do sistema Macintosh com o nome de Alex. Depois do discurso ser traduzido sonoramente para inglês foi usado o software Live para fazer modelação da voz e ser dividido em amostras sonoras. A divisão das frases do discurso é feita em duas, três, ou quatro secções, dependendo do tamanho das frases. As amostras sonoras de cada uma das secções das frases são agregados no banco de amostras respectivo de forma a serem acedidos pelo Max-MSP.

O software Max-MSP [Cycling74] é um ambiente de programação visual para música, áudio e multimédia e foi o principal software usado neste projecto. Foram criados vários *patchers* com as funções de receber os dados dos sensores, mapear os seus valores em intervalos normalizados, detectar os movimentos do corpo e escolher aleatoriamente os bancos de amostras sonoras a serem activados. A divisão das secções de frases em diferentes bancos permite reconstruções aleatórias que ainda assim mantêm ou aproximam uma construção gramatical correcta.

O som criado no Max-MSP, através do movimento do performer, é enviado para uma pista de áudio do Live [Ableton] utilizando o modo *rewire*. Nessa pista foram colocados diversos filtros que são manipulados através de controladores midi, criados no Max-MSP a partir das variáveis produzidas pelos sensores na luva. Para este projecto foram seleccionados os seguintes conjuntos de filtros: Backing Vocals 1; Tremmed Verb; Shifting Beats; Damien; Dotted Eighth Note; Feed Me Back; Fat Fire Beat; Deconstruct 1; Beat Repeat Madness 3; Grain

Delay; e Abstract-Rising Reverb. São usados simultaneamente três filtros e sempre que é pressionado um botão no fato, transfere-se para um novo conjunto de três filtros a serem usados. Desta forma o Live, através das variáveis da luva e do botão, transmitidos pelo Max-MSP, filtra o som gerado criando novas texturas sonoras.

#### 5. PERFORMANCES

A primeira apresentação pública da performance *The last words of Domenico* foi realizada no dia 16 de Julho de 2009 no contexto da festa Bang! organizada pelo LCD - Laboratório de criação digital - no espaço Plano B, no Porto. Um vídeo dessa apresentação encontra-se publicado no blog do projecto [Silva].

A performance foi também apresentada no espaço A sala no dia 24 de Outubro de 2009 (fig. 6). A Sala, é um espaço de mostra de criações performativas localizado num apartamento de habitação na baixa do Porto. Apresenta, uma vez por mês, projectos performativos de autores de diversas áreas (artes plásticas, música, dança, teatro, entre outras). Esta performance foi realizada de forma a procurar um contacto mais próximo entre o performer o público, com a intenção de se perceber se era possível existir um diálogo entre ambos. O facto é que isso não se veio a verificar, uma vez que o público, perante essa situação, manteve-se sempre numa atitude passiva. As razões para isso podem ter haver com os factos do discurso gerado pelo performer ser, sobretudo, descontinuo e absurdo, ou pela figura ser algo intimidatória.

No entanto, o principal objectivo desta performance não é a realização de um diálogo entre o performer e o público, ficando sempre essa possibilidade em aberto, mas sobretudo, o gerar de um novo discurso a partir de um discurso pré-seleccionado, com o intuito de procurar novos sentidos ao mesmo.



**Figura 6: Performance no espaço A sala.**

#### 6. AVALIAÇÃO DO FATO SENSÍVEL WIRELESS

Em entrevista realizada, em Agosto de 2009, a Joclécio Azevedo, coreógrafo e performer, e um dos directores do Núcleo de Experimentação Coreográfica, Porto (N.E.C.), procedeu-se a uma análise e avaliação do instrumento de performance dividida em dois momentos: um de

observação e outro de experimentação, de forma a se poder perceber que aspectos positivos se poderiam retirar do fato e quais as suas limitações. Por razões de espaço a entrevista não é aqui transcrita mas poderá ser consultada em [Silva09].

Num primeiro momento, a partir da observação de Azevedo do Fato Sensível Wireless, destacam-se os seguintes pontos:

- A percepção da relação entre o movimento do *performer* e o texto sonoro gerado é evidente, ainda que, por vezes, não seja perceptível qual o gesto que activa determinada amostra de áudio. Por seu lado, a importância dessa percepção só será relevante, se for essa a intenção de um projecto de aplicação deste instrumento.
- A observação desta performance sugeriu ideias e imagens associadas à robótica e ao retro-futurismo.
- As pausas são fundamentais para a percepção do discurso sonoro e, para tal, basta não existir movimento do *performer*.
- O som gerado assume mais um carácter musical e abstracto do que verbal.
- A observação do funcionamento deste instrumento gera imagens fortes, e o som gerado parece muito uniforme, podendo ser ainda melhorado nesse aspecto, mais rico na tonalidade e ritmo.
- A utilização deste fato seria mais interessante, se existisse outro fato para um segundo *performer* e assim poder-se construir um diálogo em vez de um monólogo. Artisticamente essa aproximação remeteria para uma tentativa de diálogo entre dois surdos-mudos, ou um diálogo de um casal em estado de crise, uma vez que o discurso gerado é aleatório.
- *The last words of Domenico*, é uma performance para ser realizada no meio do público, procurando dialogar com o mesmo através do gesto. A coreografia, que terá como base a improvisação, pretende estabelecer uma espécie de diálogo utópico, com um carácter algo provocativo. É por isso esperado por parte do público algum feedback para quebrar a monotonia de um monólogo. Objectivo que não é facilitado pela personagem algo inibidora sendo esse um dos desafios deste projecto.

Em segundo momento e já com base na experimentação do fato, pelo próprio Joclécio Azevedo destacam-se os seguintes pontos:

- No início tentou perceber a sensibilidade do fato e compreendeu que era necessário uma amplitude de gesto maior do que inicialmente esperado.
- Não sentiu a necessidade de controlar o gerar do discurso sonoro.
- A luva funciona muito bem, sendo o seu grau de controle elevado.
- Outros movimentos poderiam ser lidos com este fato como, o movimento da cabeça com um giroscópio.
- O fato revelou-se ser de certo modo confortável.

- À medida que se vai experimentando o fato este revela-se mais interessante, chegando mesmo a surpreender, como se tratasse de uma segunda pele do *performer*, como uma extensão do próprio corpo.

- Das melhores experiências foi sentir o silêncio nas pausas do movimento.

## 7. CONCLUSÕES

O trabalho aqui apresentado parte de um conceito artístico onde se procurava, através do movimento do corpo, encontrar novos sentidos num discurso de uma personagem – o Domenico, levando à concepção e implementação de uma interface para um *performer* – fato sensível wireless. Foi ainda feita uma confrontação com o público e com um profissional da dança de onde resultaram sugestões para actualização do fato, ideias de novas aplicações, a discussão de conceitos relacionados com a dança e performance, bem como uma reflexão sobre alguns aspectos do projecto. Ainda que na sua essência um projecto artístico, a captura de feedback dos utilizadores nas suas diferentes vertentes, veio-se revelar uma importante ferramenta de trabalho.

## 8. AGRADECIMENTOS

Os autores desejam agradecer as preciosas colaborações e suporte de: Prof. Doutora Joana Pimentel, Filipe Valpereiro, Sérgio Soares Ferreira, João Maia Silva, Joclécio Azevedo e Susana Chiocca.

## 9. REFERÊNCIAS

- [Ableton] Ableton, *Live*. <http://www.ableton.com/>
- [Bausch82] Bausch, P. 1982. *Nelken*. Excerto em: <http://www.youtube.com/watch?v=Z8wnBSclJjg&feature=related>
- [ConvenienceWare] ConvenienceWare s.d., *GhostReader*. <http://www.convenienceware.com/ghostreader.php>
- [Cycling74] Cycling74., *Max-MSP*. <http://www.cycling74.com>
- [Elliot] Elliot, C. 2003, *Transhumanism - Humanity 2.0*. <http://www.mindfully.org/Technology/2003/Transhumanist-Humanity1sep03.htm>
- [Rüdiger07] Rüdiger, Francisco. Breve história do pós-humanismo: Elementos de genealogia e criticismo. *E-Compós*; vol. 8 (2007).
- [Satomi08] Satomi, Mika, Hannah, Perner-Wilson, KOBAKANT Language game, Perfect Human. 2008. <http://www.kobakant.at/index.php?menu=2&work=5>.
- [Silva] Silva, Vitor, *The last words of Domenico*, Blog. <http://thelastwordsofdomenico.wordpress.com>
- [Silva09] Silva, Vitor 2009, *The last words of Domenico*, Tese de Mestrado, Universidade do Minho.
- [Tarkovsky83] Tarkovsky, A. 1983. *Nostalghia*. Excerto disponível em: <http://www.youtube.com/watch?v=Hm-QEjMzUp4&feature=related>

## **Posters**

Aveiro, 13, 14 e 15 de Outubro de 2010

# Metodologias de análise visual baseadas na tecnologia de Realidade Virtual: Jardim Botânico do Rio de Janeiro, Brasil e Jardim do Cerco, Portugal

**Aurélio Nogueira**

Univ. Federal do Rio de Janeiro  
Rio de Janeiro, Brasil  
aurelio.nogueira@gmail.com

**Teresa Heitor**

Instituto Superior Técnico  
Lisboa, Portugal  
teresa@civil.ist.utl.pt

**Maria Carreira**

Instituto Superior Técnico  
Lisboa, Portugal  
maria@civil.ist.utl.pt

**Gelly Rodrigues**

Instituto Superior Técnico  
Lisboa, Portugal  
gellymendes@gmail.com

**Ana Pedroso**

Instituto Superior Técnico  
Lisboa, Portugal  
luisapedroso@gmail.com

## Sumário

*Este artigo refere-se a um estudo de investigação desenvolvido no âmbito do projecto IN\_LEARNING, visando o desenvolvimento de uma ferramenta de descrição morfológica à escala urbana, através de metodologias de análise visual baseadas na tecnologia da Realidade Virtual (RV). A investigação foi desenvolvida com base em dois estudos de caso integrados no conceito de paisagem cultural, nomeadamente os jardins históricos em contextos urbanos – Jardim Botânico do Rio de Janeiro, Brasil e Jardim do Cerco em Mafra, Portugal. A descrição morfológica considera a sua dimensões históricas, geográficas e formais e é baseada na bibliografia e iconografia existente, complementada por trabalho de campo. O artigo discute a contribuição de RV para análise visual e a sua capacidade para suportar descrições espaciais. O artigo está organizado em três partes. Na primeira introduzem-se os dois estudos de caso: Jardim Botânico do Rio de Janeiro, Brasil e Jardim do Cerco em Mafra, Portugal. Na segunda apresentam-se os casos de estudo. Na terceira parte descrevem-se os procedimentos metodológicos aplicados e as principais tarefas realizadas. Finalmente, é discutido o potencial do método proposto para a descrição espacial.*

## Palavras-chave

*Meios Interactivos Electrónicos, RV, Passeio virtual, Jardim Botânico, Jardim do Cerco.*

## 1. INTRODUÇÃO

Este artigo refere-se a um trabalho de investigação realizado no contexto do projecto IN\_LEARNING, tendo como principal objecto de investigação a exploração de formas interactivas de descrição morfológica, que permitam agilizar a aquisição e a partilha de conhecimento sobre a forma urbana e funcionar simultaneamente como uma ferramenta de aprendizagem activa, i.e. como um processo simultâneo de construção do conhecimento e de descoberta, exploração e de observação.

Pretendeu-se desenvolver uma ferramenta de descrição morfológica abrangente, capaz de incorporar aspectos de natureza física e cultural, através do recurso a metodologias de análise visual baseadas na tecnologia Realidade Virtual (RV), possibilitando a criação e visualização de imagens estereoscópicas e a manipulação interactiva de modelos virtuais através do computador e facilitando a realização de experiências não convencionais em ambientes colaborativos e de partilha de informação. Esta evolução tecnológica permite conduzir a investigação no campo da descrição morfológica para novos rumos, enfa-

tizando a utilização de novas ferramentas e propiciando uma evolução no processo de descrição [Ferreira, 2004].

Para o efeito, seleccionou-se como objecto de estudo o jardim histórico em contexto urbano. Considerado como um valor paisagístico, enquadrado no conceito de paisagem cultural, a sua descrição morfológica justifica-se tanto pelo seu significado cultural, enquanto expressão de um ambiente historicamente produzido, como pela sua relevância para a leitura e qualificação da cidade. De acordo com a Carta de Florença [1981], artigo 25 deve ser estimulada a promoção património histórico. Mudur et al. [1999] reforça que existe uma necessidade absoluta de preservar e catalogar o património para um futuro estudo e memória. Pretende-se deste modo desenvolver um catálogo virtual para Web e multimédia, através de panorâmicas<sup>1</sup> de 360° baseado em dois casos de estudo: o Jardim Botânico do Rio de Janeiro (JBRJ), Brasil, e o Jardim do Cerco (JC) em Mafra, Portugal. Este sistema

<sup>1</sup> RIGG [2005] define a panorâmica como sendo uma imagem que proporciona um campo de visão maior que se pode ver.

proporciona um grande potencial para ser explorado no processo de descrição morfológica, podendo ser usado como meio de registo e de interpretação do existente, tirando partido de uma descrição tridimensional.

## 2. INTRODUÇÃO HISTÓRICA

O JBRJ teve a sua origem com a chegada da Família Real ao Brasil, em 1808. É um dos pontos turísticos e históricos mais visitados da cidade, o qual representa um instrumento de educação e consciencialização para a conservação da biodiversidade no Brasil [Koatz, 2006].

O JC situa-se entre o Palácio-Convento de Mafra e a Tapada de Mafra. Construído e traçado entre o Palácio-Convento de Mafra, a maior construção barroca do país, e a Tapada de Mafra, o maior jardim murado a nível nacional, este jardim tem o potencial único de articular estes dois valores – arquitectónico e ecológico – e juntar as duas peças da mais forte afirmação cultural da época barroca em Portugal.

## 3. METODOLOGIA

Os procedimentos metodológicos aplicados foram estruturados em cinco etapas sequenciais. (1) Na primeira etapa foram realizados trabalhos de consulta a fontes bibliográficas e iconográficas. (2) Na segunda etapa, no ano curricular de 2009/2010, realizou-se o trabalho de campo nas cidades do Rio de Janeiro e de Mafra. (3) Para a terceira etapa foi feita uma análise e selecção dos requisitos de um sistema de análise visual baseada na tecnologia RV, através de panorâmicas do JBRJ e do JC. Optou-se pela utilização de dois outros sistemas para visualização do estudo: através de PC ou um portátil com ecrã (não estéreo) on-line e off-line (multimédia). Os aplicativos serão utilizados através de uma página Web ou um multimédia gerando hipertextos, imagens e panorâmicas (possibilitando o ajuste em tempo real de qualquer eixo x, y, z). Novos recursos adicionais, como dispositivos periféricos de interface, a adição de som, odor e apresentador virtual, podem ser eventualmente adicionados. Propõem-se contudo, num futuro próximo, um projecto que contemple a concepção de uma sala de visualização móvel, com visão estéreo, a ser financiada por organização de fomento. Poderia integrar um processo de ensino/aprendizagem para um público-alvo de maior dimensão, a uma escala nacional, com a visualização dos jardins, através de um multimédia com imagens, hipertextos entre outros. (4) A quarta etapa teve como objectivo o registo fotográfico digital dos diversos locais do JBRJ e JC e a elaboração das imagens panorâmicas. (5) Na quinta etapa foi elaborado o passeio virtual interactivo, através de panorâmicas executadas na fase anterior acrescidas de hipertextos e fotos interactivas. O passeio virtual está estruturado com base numa planta, onde existem pequenos ícones coloridos. Ao clicar em cada ícone ver-se-á a panorâmica correspondente a esse local (planificada e interactiva), bem como texto explicativo com toda informação histórica relativa a esse espaço. É possível passear virtualmente no jardim, clicando nos pontos acima referidos para visualização das imagens e informações correspondentes.

## 4. O POTENCIAL DO MÉTODO PROPOSTO PARA A DESCRIÇÃO MORFOLÓGICA

Considerando que a função do património é ser construtiva [Choay 2001] torna-se fundamental não só perpetuar os testemunhos do passado como dialogar com esse passado, através da sua apropriação e releitura, permitindo a sua fruição. A RV, como nova forma de descrição morfológica, não irá substituir as tecnologias já existentes, mas sim, complementá-las [Campos e Sampaio 2005], constituindo um instrumento de grande eficácia. A tecnologia utilizada permite obter visualizações tridimensionais e, simultaneamente, integrar uma grande variedade de dados e imagens num ambiente de inclusão digital e cultural extrapolando os limites espaço-tempo, possibilitando a exploração virtual de lugares e agindo como uma espécie de “máquina do tempo” [Dainese et al 2003].

A ilustração da RV promove ainda uma análise da realidade visualizada sob diferentes ângulos com sentimento de imersão virtual, permitindo a visualização e exploração de lugares de difícil acesso, de forma a compreender a espacialidade do caso de estudo, bem como todas as características morfológicas do espaço. Ao agilizar a aquisição e a partilha de conhecimento sobre a forma urbana, o uso de RV contribui para o resgate da sua memória, funcionando simultaneamente como uma ferramenta de aprendizagem activa, possibilitando a partilha e discussão de ideias, de forma rápida e inovadora, com perfeita visualização e compreensão dos espaços.

## 5. REFERÊNCIAS

- CAMPOS, A., SAMPAIO, F. (2005) “Uma Aplicação de Realidade Virtual não Imersiva no Ensino de Astronomia”. XVI Simpósio Brasileiro de Informática na Educação, Juiz de Fora-MG.
- CHOAY, F. (2001) “The Invention of the Historic Monument”, Cambridge: Cambridge University Press
- DAINESE, C., et al. (2003) “Sistema de Realidade Aumentada para Desenvolvimento Cognitivo da Criança Surda”. In: Anais do 6th SBC Symposium on Virtual Reality, Ribeirão Preto, Brasil, 273-282.
- FERREIRA, L. F. (2004). “Usando Objectos educacionais baseados em Realidade Virtual em ambientes de apoio a construção de conhecimento e aprendizagem de técnicas videocirúrgicas”. CINTED Novas Tecnologias na Educação. Vol. 2, Nº 1, UFRG.
- KOATZ, G. (2006). “Condições de vida no bairro do Jardim Botânico e trecho da Lagoa adjacente”. Tese Doutorado, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- MUDUR, et al (1999) “3-Dimensional Documentation of “Complex Heritage Structures” Available in <http://westwood.fortunecity.com/karan/133/paper.htm> Acedido em Abril 2010.
- RIGG, J. The Guide “To Panoramas and Panoramic Photography”. <<http://www.pinefarm.demo.co.uk/james/panoquide/>>, Acedido em Março 2005.



# Áudio Descrição na Televisão Digital Terrestre (TDT)

Rita Oliveira  
DeCA, Universidade de Aveiro  
Aveiro, Portugal  
ritaoliveira@ua.pt

Jorge Ferraz de Abreu  
DeCA, Universidade de Aveiro /  
Cetac.media  
Aveiro, Portugal  
jfa@ua.pt

Ana Margarida Almeida  
DeCA, Universidade de Aveiro /  
Cetac.media  
Aveiro, Portugal  
marga@ua.pt

## Sumário

*A Televisão Digital Terrestre (TDT), aliada à utilização de Set-Top Boxes (STB), permite aos telespectadores a utilização de funcionalidades e serviços avançados que o sistema de teledifusão convencional não consegue suportar. Para a maioria das pessoas a experiência de utilização televisiva na TDT é fácil e simples, pois o uso das novas funcionalidades e serviços não requer conhecimentos tecnológicos específicos. No entanto, Portugal possui um número expressivo de pessoas com deficiência visual, que não podem beneficiar em absoluto desta nova tecnologia. Este artigo apresenta os objectivos e a metodologia que suportam um projecto de investigação ligado à problemática do Design Universal aplicado à TDT, no qual é destacada a possibilidade da integração do meio complementar de comunicação televisivo da áudio descrição em contexto interactivo.*

## Palavras-chave

*Design Universal, Televisão Interactiva, Televisão Digital Terrestre, Áudio Descrição*

## 1. INTRODUÇÃO

Os serviços interactivos actualmente fornecidos pelos operadores de televisão digital envolvem uma forte componente visual, apresentando dificuldades acrescidas de interacção por parte de utilizadores com deficiência visual (UDV). Como resultado, para além dos problemas na interpretação do conteúdo televisivo, estes utilizadores possuem dificuldades acrescidas no acesso a este tipo de serviços. Deste modo, torna-se importante a aplicação de princípios de acessibilidade, usabilidade e design orientados ao desenvolvimento de sistemas de interacção televisivos acessíveis.

Com a introdução da TDT em Portugal, segundo a Lei da Televisão n.º 27/2007 de 30 de Julho [AR07], terá de ser assegurado o complemento da emissão de uma parte da programação televisiva através de meios auxiliares de comunicação (áudio descrição, legendagem e interpretação gestual), tal como acontece actualmente com a transmissão analógica. As STB associadas à TDT, atendendo à sua crescente capacidade computacional e possibilidade de ligação à Internet, poderão facilitar a implementação destes meios auxiliares de comunicação.

Neste cenário, o projecto de investigação que aqui se apresenta tem como principal objectivo a conceptualização, prototipagem e validação do modelo de interacção de um serviço de áudio descrição disponibilizado pela TDT, que responda às necessidades dos UDV.

## 2. OBJECTIVOS DE INVESTIGAÇÃO

A presente investigação integra três etapas interdependentes entre si. Na primeira etapa pretende-se efectuar o

diagnóstico das dificuldades e necessidades de UDV enquanto consumidores de serviços de áudio descrição interactivos. Já na segunda etapa pretende-se desenvolver um protótipo de um sistema de áudio descrição interactivo suportado pela TDT, que integre novas funcionalidades e que se baseie em princípios de design universal. Por fim, a terceira etapa contempla o teste e a avaliação do modelo de interacção prototipado, por parte de UDV.

Deste modo, esta investigação tem como finalidades:

- Contribuir para o desenvolvimento de estratégias de design universal nos modelos de interacção para TV, nomeadamente na TDT.
- Promover a inclusão digital dos consumidores televisivos com deficiência visual através da utilização de um serviço interactivo de áudio descrição integrado na TDT.

Para facilitar a operacionalização da investigação, especificaram-se os seguintes objectivos:

- Identificar e caracterizar as especificidades e necessidades dos UDV, enquanto consumidores de conteúdos televisivos.
- Analisar o design de interacção de serviços de áudio descrição interactivos já existentes.
- Analisar e compreender o contexto funcional e tecnológico da TDT e respectivas STB.
- Identificar os princípios de design apropriados ao desenvolvimento de um sistema interactivo de áudio descrição que responda às necessidades dos UDV e,

numa lógica de design universal, possa acomodar diferentes cenários de utilização e públicos-alvo.

- Conceptualizar e prototipar um modelo de interacção de um serviço de áudio descrição disponibilizado pela TDT, que integre novas funcionalidades e seja baseado nos princípios de design identificados.
- Testar e avaliar o modelo de interacção junto dos públicos-alvo.

### 3. METODOLOGIA DE INVESTIGAÇÃO

Com o objectivo de operacionalizar a investigação, combinam-se dois métodos qualitativos que correspondem à primeira e terceira etapas identificadas anteriormente, sendo elas a Teoria Fundamentada nos Dados (*Grounded Theory*) e o Estudo de Caso, optando-se assim por uma triangulação de métodos. Segundo Carmo e Ferreira [Carmo&Ferreira08], esta integração metodológica promove uma melhor ilustração e compreensão do fenómeno em estudo, permitindo estabelecer relações entre os resultados obtidos.

A Teoria Fundamentada nos Dados consiste num conjunto de procedimentos metodológicos que tem como principal objectivo a análise de dados, descritos e organizados numa sequência evolutiva que procura a integração desses mesmos dados [Glaser&Strauss67]. Assim, visa a recolha e a análise sistemática dos dados, orientando os investigadores através de um método indutivo de criação de conhecimento. Para além destas características, a comparação constante é o princípio central desta teoria, na qual a interpretação e conceptualização dos dados têm um papel fundamental. A justificação da escolha deste método relaciona-se com a heterogeneidade do público-alvo e com o carácter inovador e simultaneamente incipiente dos serviços interactivos da TDT, pois pretende-se obter dados antes de se avançar para a construção de interpretações teóricas. Neste sentido, as noções teóricas acerca do fenómeno em estudo evoluem durante a investigação, ocorrendo através da relação dinâmica e contínua entre a análise e a recolha de dados. A primeira etapa do trabalho de investigação tem por base este método, para que na sua conclusão se obtenham dados expressivos e suficientes para determinar as dificuldades e necessidades dos UDV enquanto consumidores de serviços de áudio descrição interactivos. No que diz respeito à terceira etapa de investigação, esta encontra-se dependente dos resultados obtidos na primeira etapa. Se os resultados justificarem a conceptualização e prototipagem de um novo modelo de interacção, é desencadeada a segunda etapa de investigação. Após a finalização desta etapa, para que seja possível perceber se o modelo de interacção subjacente ao sistema prototipado responde às necessidades dos UDV é desencadeada a terceira etapa de investigação. O Estudo de Caso é o método que suporta esta etapa de investigação, pois assume-se como uma pesquisa particular e singular que procura descobrir o que há de mais essencial e específico na situação em análise, possuindo como objecto de estudo uma entidade bem definida [Pardal&Correia95]. Para além disso, baseia-se numa investigação empírica que analisa um fenómeno actual

em contexto real, nomeadamente quando os limites entre o fenómeno e o contexto não estão claramente definidos [Yin05]. No caso específico desta investigação, o fenómeno em estudo centra-se na acessibilidade e usabilidade do modelo de interacção prototipado, detendo como objecto de estudo os sujeitos com deficiência visual que irão testar e avaliar esse modelo. No que diz respeito ao tipo de estudo de caso, este classifica-se como um estudo de caso prático [Pardal&Correia95], pois possui uma natureza utilitária, tendo como objectivo principal a realização do teste e avaliação do modelo de interacção prototipado junto do público-alvo a que é dirigido. Por outro lado, segundo a classificação de Yin [Yin05], este estudo de caso é do tipo explanatório, já que a experimentação surge como a estratégia de pesquisa eleita e a situação em estudo possui ligações operacionais que necessitam de ser explicadas a partir de uma teoria.

### 4. COMENTÁRIOS FINAIS

O projecto de investigação apresentado possui como pressupostos o estudo de novas estratégias para a integração da pessoa com deficiência visual no plano da Televisão Interactiva, como também a criação de um modelo de interacção de um serviço de áudio descrição que responda às especificidades e necessidades deste tipo de utilizadores. “*O desenvolvimento tecnológico ao serviço de uma cidadania moderna e de progresso exige uma sociedade da informação inclusiva onde o conhecimento é um valor ético, social, cultural e económico fundamental que promove a criação de riqueza e emprego, a qualidade de vida e o desenvolvimento social*” [UMIC10]. A partir desta afirmação, a relevância social desta investigação é clara, na medida em que surge como um desafio para políticas de regulamentação, em termos do acesso universal, da promoção da diversidade e do pluralismo e da protecção dos interesses dos consumidores televisivos com necessidades especiais.

### 5. REFERÊNCIAS

- [AR07] Assembleia da República (AR): *Lei da Televisão – Lei n.º 27/2007. D.R. n.º 145, Série I de 2007-07-30, Diário da República Electrónico, 2007, [url]: <http://dre.pt/pdf1sdip/2007/07/14500/0484704865.pdf> (12-05-2010).*
- [Carmo&Ferreira08] Carmo, H. e Ferreira, M. (2008): *Metodologia da Investigação – Guia para Auto-aprendizagem*, 2ª Edição, Universidade Aberta: Lisboa.
- [Glaser&Strauss67] Glaser, B. e Strauss, A. (1967): *The discovery of Grounded Theory: strategies for qualitative research*, Aldine: Chicago.
- [Pardal&Correia95] Pardal, L. e Correia, E. (1995): *Métodos e técnicas de investigação social*, Areal Editores: Porto.
- [UMIC10] Agência para a Sociedade do Conhecimento (UMIC): *UMIC – Início*, Portal UMIC, 2010, [url]: <http://www.umic.pt/> (25-05-2010).
- [Yin05] Yin, R. (2005): *Estudo de caso : planeamento e métodos*, 3ª Edição, Bookman: Porto Alegre.

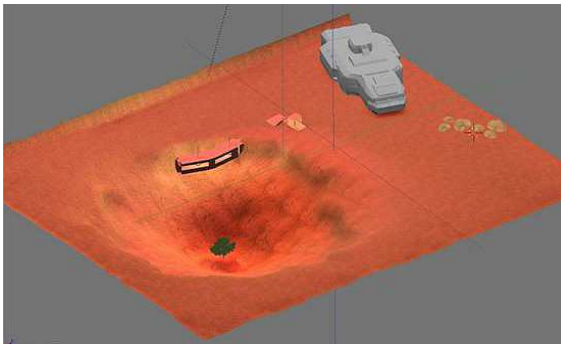
# Tree on Mars, an Immersive Virtual Reality Experience

Paulo Ricardo Duarte  
MEI-M ISCTE-IUL – Lisboa, Portugal  
Biodroid Productions, SA  
pauloduartewalac@gmail.com

Miguel Sales Dias  
ISCTE-IUL – Lisboa, Portugal  
MLDC, Microsoft Language Development  
Center – Porto Salvo  
miguel.dias@microsoft.com

## 1. INTRODUCTION

Tree on Mars is an Immersive Reality project, with the following story: the viewer is an astronaut who travels to Mars aiming to examine a tree found in the center of a crater. The project is to be viewed in the CaveH platform located at Centro de Ciência Viva at Lousal, and was developed in the context of the 3D programming curriculum of the Masters of Informatics Engineering - Multimedia of ISCTE-IUL in Lisbon. By definition, a Cave is an Immersive Virtual Reality (VR) system, which projects stereoscopic images with real scale dimensions [1]. In the CaveH case, the set-up is of a parallelepiped form and each face is a planar projection image, where the 3D effect comes from its stereoscopic projection.



**Figure 1: Tree on Mars landscape**

The scenario of Mars has one base facing the tree, a spaceship, a rock and an Explore Mars Rover, as shown in Figure 1. In this poster we will highlight the tools and 3D authoring techniques appropriate for real-time 3D rendering, in large scale immersive virtual reality settings. We will illustrate the various stages of development and the achieved results.

## 2. AUTHORIZING CONTENT FOR CAVEH

The CaveH content authoring and development process goes through several phases. The process starts by modeling the 3D objects in an authoring tool such as Blender [2]. Then, the content is exported to the Open Scene Graph (OSG) [3] format, to be inserted in the respective folders of GTKRadiant [4], the free game engine used in the CaveH content authoring environment. GtkRadiant is able to author and run the 3D scenario, where the OSG files are included, with many authoring possibilities, like setting Newton physics simulation, including collision detection and response between 3D objects, key-frame forward kinematic animation of characters and virtual camera and visual effects properties, such as sky-dome and particle

systems and others. After the authoring is completed in GtkRadiant, CaveHSpawner compiles the map, 3D models and their properties in an OSG scene graph and, upon error free compilation, generates an octree-based volumetric organization of the scene-graph, with Level of Detail nodes for the more complex 3D models, ready for real-time rendering. Finally, the user is able to run the CaveH runtime viewer application (in a single projection system such as a PC or a multi-projection setting, such as a CAVE), with the compiled map. OpenSceneGraph is the adopted open source 3D graphics scene graph and toolkit, specially developed for real-time rendering or large and complex 3D scene graphs. The OSG models can be generated in two different ways:

1. By producing OSG objects from scratch with the C++ programming language using, for example, Microsoft Visual Studio C++ development environment.
2. By exporting 3D models from Blender with OSG Exporter Script [2], based on the Python language. This script transforms Blender models into OSG objects.

The second method was the chosen option. All models were created in Blender and then exported to OSG objects. The content was developed on a computer with the following specs: 1.60GHZ 2x CPU, 1 GB Ram, and NVIDIA GeForce 7900 GT/GTO 512 MB graphic card. As tools, we have used the following one: Blender, Photoshop, GtkRadiant 1.5 and CaveHSpawner, under MS Windows.

### 2.1 SETTING-UP BLENDER FOR REAL-TIME

Blender is an open source 3D content creation tool that uses Python as its scripting language. Blender was primarily designed for making non real-time scenes, rendered with a ray tracer. It has a different way of processing material properties and lights, which doesn't quite fit with real-time formats, like Collada and OSG own format. For real-time rendering we need to understand how to set resources such as: materials, normals, transforms and reference frames:

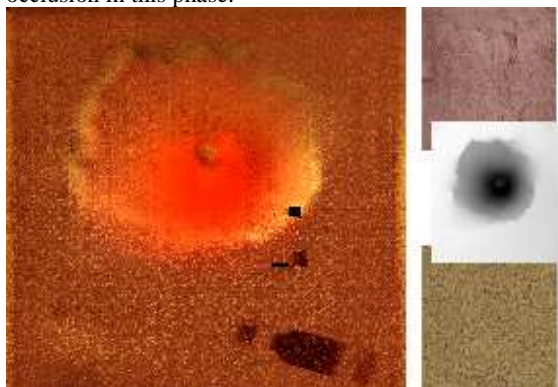
- Materials: applying materials for real-time use is different than for non real-time applications. It is better to use one material per model; this stands for textures as well. Additionally, the texture has to be mapped onto the object using UV coordinates.
- Normals: all normals should face outwards. If not, the polygons will not be shown in the 3D view.
- Transforms and reference frames: without care, an object will spin around its local reference frame. Values can be set for Location, Rotation and Scale. If Rotation gets non-zero values, this object is going to

rotate the way it is expected at runtime. All of this can be adjust and fixed by using Apply Scale and Rotation in Blender.

## 2.2 REAL-TIME TECHNIQUES

To map correctly the geometry we have used some techniques normally adopted in real-time productions. In the Mars landscape we have adopted a common technique which raises soil with a Height Map. A Height Map [5] is a grayscale image, where pure white represents the highest points and pure black the lowest points in the set. The Height Map is then applied onto a subdivided plane, where the ground is raised or lowered by controlling the grayscale vertices.

As for texturing, this is a relatively complex task. First comes the UV mapping of the 3D models. Each object produces shadows and reflections. These are calculated by baking everything – the texture, shadows, shading and ambient occlusion – in one simple UV texture. It is critical to set appropriately the light and the ambient occlusion in this phase.



**Figure 2:** The left image is the result of the Stencil technique used in the right image.

For mapping the Mars soil we have used a Stencil technique [5]. This method uses 3 textures to blend 2 via the grayscale of another, as shown in Figure 2. The upper texture blends dark areas of grayscale Stencil and the lower texture blends white areas (the blend method used is overlay). The final texture (left in Figure 2) shows some after effects made on Photoshop. This grayscale image was also used for the Height Map. There are specific settings in the textures that must be considered to obtain the maximum quality of the OSG files [3]. To have alpha channel working well on OSG objects, the texture must have 32-Bit color and has to be saved in a format that holds the alpha channel, like .png or .dds. The .dds file was the chosen format for two reasons: first, this is the format for videogames proposed by Nvidia [6] and second, this format can save the same detail in less disk space than .png.

## 2.3 GTK RADIANT

After having exported and saved all 3DS and OSG files into their own GtkRadiant folders, the next step was the creation of the scene. First, all the collider parameters were interactively defined for each 3D object and, whenever necessary, Newton physics was enabled. Next, we have included a representation of a first-person view, where it is possible to modify a number of human scales,

motion parameters and fog options. We have also included a light and a skybox, which simulates the environment with dynamic clouds. A particle system allowed us the insertion of flying leaves over the tree. Lastly, we have used a trigger on the first-person view location to load a movie in the beginning of the interaction and a sound listener to set on the environmental music. Finally, GtkRadiant can save a map file (with all the authoring changes made), that will be loaded in another tool, the CaveSpawner, for compilation and run-time execution.

## 3. CONCLUSIONS

The development of 3D content for real-time rendering is far different than for traditional ray-tracing rendering. For real-time uses, there are a number of techniques that requires different modeling skills and some degree of originality. In real-time 3D rendering, we have to simulate shadows and reflections by using texture baking, we need to be aware of normals and polygons counts of our 3D models and, last but not least, we need to have in mind the level design related with the storyboard. In real time, there is not a single camera to guide us by the desired key-frame spots. Here the camera is the player that has the possibility to see what he/she wants, so the level needs a trustworthy environment and a world limit must be imposed. The CaveH multi-projection system is a brand new environment in Portugal for Immersive Virtual Reality. A field of view of 180° provides more credible immersions and makes the player believe that he or she is in a real scale landscape. In this poster, we have described an immersive virtual reality experience carried in the CaveH of Lousal, highlighting specially the content authoring techniques that are more appropriate for real-time rendering. The experience allowed the learning of some concepts about Scene Graph development, the correct parameterization of modeling techniques for real-time image synthesis and the integration of some classical animation techniques that, in spite of being largely used in the 90's for computer graphics, are still in use today, namely in gaming, like action triggering and key-frame animation. Perhaps very soon the entertainment will be projected on entire room walls with the same classic 90's video game techniques.

## 4. REFERENCES

- [1] Bastos, P., Dias, M. S., "Experiência de Realidade Virtual Imersiva no Ambiente CaveHollowspace do Lousal", Proceedings of Interação 2008 – 3ª Conferência Interação Pessoa-Máquina, Universidade de Évora, 15 – 17 Oct 2008.
- [2] Blender. <<http://www.blender.org/>> and OSG Script. <<http://vterrain.org/Doc/Blender/>>
- [3] Martz, P., *OpenSceneGraph - A Quick Introduction to the Cross-Platform OSG API*. Computer Graphics Systems Development Corp., 2007.
- [4] GtkRadiant. <<http://qeradiant.com/cgi-bin/trac.cgi>>
- [5] Luke A., *3D Game Textures*, 2ª Edition. Focal Press, 2009, 296-319.
- [6] NVIDIA DDS Plug-in for Adobe Photoshop. <<http://developer.nvidia.com>>



# Elaboração de Conteúdos 3D para utilização na CaveHollowspace do Lousal

Luís Miguel S Ponciano  
MEI-M ISCTE/IUL - Lisboa  
lmspo@iscte.pt

Miguel Sales Dias  
ISCTE-IUL – Lisboa  
MLDC, Microsoft Language Development  
Center – Porto Salvo  
miguel.dias@microsoft.com

---

## Sumário

*Este poster descreve as diversas etapas da elaboração de conteúdo 3D para utilização no ambiente de realidade virtual imersivo instalado no Centro de Ciência Viva do Lousal.*

## Palavras-chave

*Realidade Virtual Imersiva, Modelação 3D, Autoria 3D.*

---

## 1. INTRODUÇÃO

Os sistemas de realidade virtual imersiva são a tecnologia que permitem ao ser humano ter a sensação e a ilusão de estar num mundo virtual criado digitalmente [1]. Na conferência SIGGRAPH de 1992 foi apresentado o primeiro ambiente virtual fechado CAVE (Cave Automatic Virtual Environment), desenvolvido na Universidade do Illinois-Chicago's [2]. Em 2007 é criado o primeiro ambiente de realidade virtual imersiva em Portugal com uma capacidade para grupos até 14 pessoas, no Centro de Ciência viva localizado nas antigas minas do Lousal [3].

Este *poster* descreve as etapas necessárias para a elaboração de conteúdos 3D, para o ambiente imersivo de realidade virtual instalado no Centro de Ciência Viva do Lousal, mas cuja aplicação se pode generalizar a qualquer sistema de multi-projecção que execute a plataforma de *middleware* e de gestão de conteúdos CaveH.

## 2. ARQUITECTURA

A arquitectura do ambiente imersivo de realidade virtual instalado no Lousal é composta por seis planos de projecção numa topologia em U, com as dimensões 5.6m x 2.7m x 3.4m que permite um campo de visão superior a 180° (ver Figura 1), estereoscopia através de óculos passivos INFITEC, seguimento do utilizador por infra-vermelhos, um interface multimodal através do *Wiimote & Nunchuk* [4] e suporte a reconhecimento de fala e um sistema de som de alta qualidade com distribuição 7.1 e *subwoofer*.

A camada lógica assenta numa plataforma Microsoft Windows. Ao nível do sistema gráfico é utilizada plataforma de *software* aberto OpenSceneGraph [4] para gerir o grafo de cena 3D que contém diversos módulos incorporados através de um sistema de *plugins*, tais como

simulação de sistemas de partículas, suporte avançado para o carregamento de diversos tipos de modelos geométricos (3D Studio, Lightwave, Wavefront OBJ, COLLADA, VRML 1.0, etc) bem como imagens (PNG, JPEG, TGA, TIFF, GIF, BMP, etc).

O módulo MX-Toolkit [5] foi adaptada para gerir a interacção multimodal com o utilizador e o seu seguimento no espaço de trabalho da CaveH, estando integrada no respectivo *middleware*.

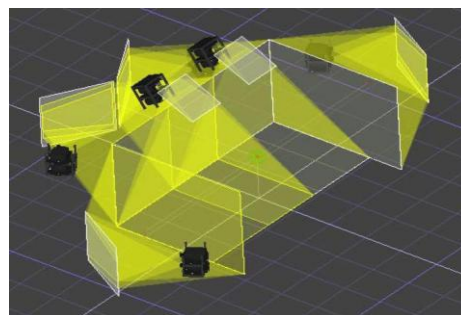


Figura 1 – Disposição dos planos de projecção na Cave-Hollowspace do Lousal.

## 3. CRIAÇÃO DE CONTEÚDOS

Para a elaboração dos conteúdos, foi utilizado um PC com 2Gb de *Ram* e uma placa gráfica NVIDIA GeForce GO 7400 128MB, com Windows XP. De seguida, descreve-se o ambiente específico de modelação e autoria.

### 3.1 Modelação 3D

Para a criação dos modelos 3D recorremos à ferramenta Blender [6] que exportou os modelos em formato 3DS, posteriormente convertidos para OSG através da aplicação *osgconv.exe* do OpenSceneGraph (ver Figura

2a). Alternativamente, com recurso à linguagem de programação C++, em ambiente Microsoft Visual Studio 2003 e com as bibliotecas do OpenSceneGraph, foram criados vários objectos com o formato OSG, dos quais destacamos o efeito do jacto de lava que foi conseguido com base no sistema de partículas do OpenSceneGraph [4] (ver Figura 2b).

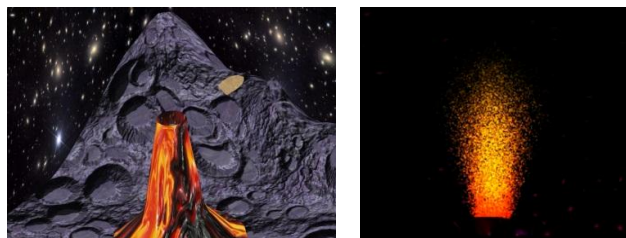


Figura 2 – (a) Cenário principal, (b) de jacto de lava.

### 3.2 Autoria do Ambiente Virtual 3D

A ferramenta de autoria utilizada é o GtkRadiant [7], que permite dispor os objectos no cenário 3D e criar interacções entre o “jogador” que controla a câmara virtual e os referidos objectos da cena, através de sensores de distância ou temporais (*triggers*), que permitem disparar uma série de acções, conforme se pode visualizar na Figura 3.

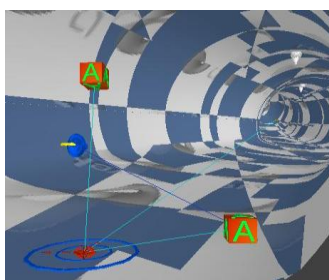


Figura 3. –Vista de objectos e *triggers* no GTK.

Além dos objectos modelados, podemos adicionar a esta plataforma de autoria, diversos tipos de fonte de luz, ficheiros de áudio e animações cinemáticas desenvolvidas numa outra ferramenta disponibilizada pela plataforma, o CaveH *Fly Editor*. Todo o processo final passa pela ferramenta CaveH *Spawner*, por forma a compilar (integrar os objectos e a cena num grafo de cena OpenSceneGraph e organizar volumetricamente a cena numa árvore de octantes), e executar a experiência de realidade virtual, num ambiente não imersivo (PC) ou imersivo (um sistema de multi-projecção como a CaveH ou outro). Na figura 4 podemos visualizar uma vista do início da experiência de realidade virtual produzida.

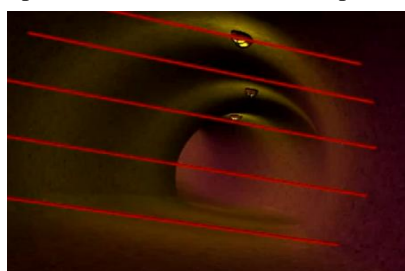


Figura 4. – Vista cenário, no ponto de início da experiência de realidade virtual imersiva.

## 4. CONCLUSÃO

Neste *poster* foi sucintamente descrita a infra-estrutura de realidade virtual imersiva CaveH, instalada no Centro de Ciência Viva do Lousal. No contexto da criação de conteúdos para esta infra-estrutura, foram abordados dois processos de modelação, um através de ferramentas existentes na comunidade que geram objectos estáticos e um outro processo da criação de objectos, neste caso em C++ e com o recurso as bibliotecas do OpenSceneGraph [4]. Esta última técnica permite a criação de objectos com algum comportamento específico, o que permitiu ao projecto colocar na cena, objectos com animação e comportamentos autónomos. Ao nível da autoria foram descritas a disposição e a interacção dos objectos com o cenário. Este projecto foi elaborado no âmbito do trabalho prático para a unidade curricular: Programação 3D, do Mestrado de Eng. Informática - Multimédia do ISCTE/IUL[8]. Conclui-se que, ao nível da criação de objectos, a plataforma de autoria é bastante flexível, pois permite-nos incluir uma vasta gama de formatos 3D. O sistema de suporte a essa autoria, o editor do GTKRadiant, permite dispor com facilidade todos objectos na cena e definir o nível de interacção, de uma forma acessível e precisa. As tarefas de compilação e execução do conteúdo, quer em ambiente PC quer em ambiente CAVE, ficam a cargo das ferramentas CaveH o que facilita todo este processo. Por outro lado, o recurso às bibliotecas do Openscenegraph [4], uma plataforma aberta e sempre em expansão por uma comunidade académica vibrante, permite a introdução das inovações provenientes da comunidade científica.

## 5. REFERÊNCIAS

- [1] Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola , Ivan Poupyrev, 3D User Interfaces - Theory and Practice(Addison-Wesley) 2004.
- [2] Cruz-Neira, Sandin, D.J., DeFanti, T., “Surround-screen projection-based virtual reality”, Laboratory (EVL) at the University of Illinois at Chicago, SIGGRAPH 92, 1992.
- [3] Dias, M. S. et al., “CaveHollowspace do Lousal – Princípios Teóricos e Desenvolvimento”, Curso Curto, 15º Encontro Português de Computação Gráfica, Microsoft, Porto Salvo, Out 2007.
- [4] OpenSceneGraph 2.0.8, <<http://www.openscenegraph.org>>, acesso em 20.Jan.2010
- [5] Dias, J., M., S., Santos, P., Monteiro, L., Silvestre, R., Bastos, R., “Developing and Authoring Mixed Reality with MX Toolkit”, ART03, The Second IEEE International Augmented Reality Toolkit Workshop, Tokyo, Japan, 6 October 2003.
- [6] Blender 2.49b, <<http://www.blender.org>>, acesso em 20.Jan.2010
- [7] GTKRadiant, <<http://www.qeradiant.com>>, acesso em 20.Jan.2010
- [8] ISCTE/IUL, Disponível em <<http://www.iscte.pt>>, acesso em 20.Jan.2010.



# O Tacto e a Interacção Móvel

João Benedito   Tiago Guerreiro   Hugo Nicolau   Joaquim Jorge   Daniel Gonçalves  
 Instituto Superior Técnico / INESC-ID  
 Rua Alves Redol, 9, 100-029, Lisboa  
 jpmlb@ist.utl.pt.com, {tjvg,hman,jaj,djvg}@vimmi.inesc-id.pt

## Sumário

*Os dispositivos móveis são normalmente desenhados para utilizadores sem qualquer tipo de deficiência. Consequentemente, o retorno táctil é muitas vezes negligenciado em detrimento de dispositivos esteticamente atractivos. Mais, as soluções de acessibilidade são normalmente estereotipadas, encarando as deficiências através de uma perspectiva limitada. Em particular, os leitores de ecrã são usados por utilizadores cegos como forma de ultrapassar a incapacidade de receber retorno do dispositivo. Porém, estas soluções apenas resolvem alguns problemas existentes. Na cegueira, outras capacidades ganham uma maior relevância. Acima de tudo, o tacto desempenha um papel essencial quando se interage com teclados físicos. Para maximizar o desempenho destes utilizadores, é necessário ter um conhecimento mais aprofundado das suas capacidades. Neste trabalho propomos uma aproximação de compatibilidade utilizador-produto, tentando correlacionar a sensibilidade táctil dos utilizadores e exigências dos teclados, permitindo a criação de interfaces através de um desenho informado.*

## Palavras-chave

*Sensibilidade táctil, Cegos, Acessibilidade Móvel, Avaliação.*

## 1. INTRODUÇÃO

Os telemóveis estão cada vez mais presentes no nosso dia-a-dia, tornando-se muitas das vezes ferramentas indispensáveis. Eles oferecem-nos a capacidade de comunicarmos quando e com quem quisermos, independentemente de onde estamos. São responsáveis por um grande aumento da comunicação e, consequentemente, uma ferramenta poderosa de inclusão social. No entanto, o progresso tecnológico levou-nos ao desenvolvimento de dispositivos pequenos e atraentes, apresentando diferentes exigências aos utilizadores. Em particular, uma pessoa cega enfrenta vários desafios ao interagir com dispositivos móveis. Devido à perda de visão esta população depende principalmente da sensibilidade táctil para manipular os seus aparelhos. Ainda assim, este é um sentido que tem vindo a ser renegado na interacção móvel, em especial aquando da ausência de visão. Para além disso 82% dos cegos têm mais de 50 anos, resultando numa diminuição exponencial do tacto [WHO07].

As soluções móveis actuais limitam-se à utilização de leitores de ecrã, substituindo o retorno visual pela sua representação auditiva. No entanto, estes estão restringidos às informações contidas no ecrã. Assim, existe um esforço não só para decorar a disposição das teclas no teclado, mas também para as distinguir, dando uma enorme relevância ao tacto aquando da ausência de visão. Apesar de todos estes factores, não se sabe qual o dispositivo mais adequado para uma pessoa em particular.

O nosso objectivo é preencher esta lacuna e relacionar as capacidades tácteis dos utilizadores com as características e exigências dos dispositivos. Neste trabalho apresentamos a nossa abordagem, bem como a avaliação preliminar das exigências dos teclados, mostrando que estas podem influenciar bastante o desempenho dos utilizadores.

## 2. A RELEVÂNCIA DO TACTO

A teoria *capability-demand* baseia-se nos conceitos de capacidades dos utilizadores e exigências dos produtos, tendo como objectivo analisar a compatibilidade utilizador-produto [Persad07]. Neste trabalho vamos nos concentrar numa capacidade em particular, o tacto, e nas exigências que lhe dizem respeito.

Quanto às capacidades do utilizador, a sensibilidade táctil será avaliada em termos de acuidade espacial e sensibilidade à pressão. Quanto às exigências do produto, podemos categorizar os teclados em termos de tamanho, espaçamento, relevo, material e etiqueta das suas teclas. Com ambas as dimensões quantificadas, vamos ser capazes de identificar relações entre elas e visar uma correspondência entre as capacidades dos utilizadores e as exigências dos produtos. Assim, poderemos criar ferramentas de suporte ao desenho, mostrando quais as opções mais eficientes e inclusivas.

### 2.1 Sensibilidade táctil

Iremos focar na sensibilidade à pressão e na acuidade espacial, visto estas serem as capacidades envolvidas quando se interage com teclados. Para além disso,

incluímos um teste funcional que consiste na identificação do número de pontos de caracteres Braille, impressos aleatoriamente numa folha de papel. A sensibilidade tátil vai ser medida no polegar, apontador e dedo médio, pois estes são os dedos normalmente utilizados, de acordo com questionários preliminares já realizados.

## 2.2 Exigências dos produtos

Cada telemóvel irá ser categorizado de acordo com um conjunto de dimensões que o caracterizam, criando assim um mapa de desenho de exigências dos teclados:

**Tamanho:** teclas demasiado pequenas são mais difíceis de identificar.

**Espaçamento:** quando o espaçamento entre teclas é demasiado pequeno é mais difícil distingui-las.

**Altura:** é mais difícil distinguir uma tecla com pouca altura relativamente à base do telemóvel.

**Material:** se o material das teclas e da base do telemóvel for contrastante é mais fácil distingui-los.

**Etiqueta:** a etiqueta deve ter uma altura e número suficientes para que seja facilmente identificável.

Considerando a teoria de *capability-demand*, para a experiência do utilizador ser bem sucedida, as exigências do produto têm de ser menores que as capacidades do utilizador. É por isso necessário identificar as exigências dos teclados relativamente à sensibilidade tátil dos utilizadores. Para encontrar estas relações, além da avaliação às suas capacidades tácteis, iremos avaliar os seus níveis de eficácia e eficiência na realização de primitivas de interacção (e.g. aquisição de teclas) em diferentes telemóveis. Desta forma vai ser possível encontrar relações entre a sensibilidade tátil dos utilizadores e as características físicas dos teclados.

## 2.3 Prever a relação utilizador-dispositivo

Nesta fase, vamos confirmar se, de facto: os utilizadores têm diferentes níveis tácteis; se estes afectam a interacção; quais as avaliações mais relevantes; quais as características dos teclados que têm maior impacto; e como ambas as dimensões se relacionam. Assim, poderemos criar um modelo preditivo, utilizador-dispositivo, que para cada utilizador sugere qual o teclado móvel mais adequado às suas capacidades tácteis. Seremos também capazes de prever quais os dispositivos mais abrangentes para uma população alvo. Este modelo será avaliado com alguns dispositivos disponíveis no mercado e junto dos utilizadores. Um modelo teórico deste tipo apresenta várias vantagens. Por exemplo, uma empresa de desenho pode usá-la em fases iniciais do projecto para tomar decisões, evitando problemas de acessibilidade em etapas avançadas de produção. Para além disso, permite avaliar a inclusão de um dispositivo sem recurso a avaliações empíricas, que têm um elevado custo. Para os utilizadores será útil perceber qual o dispositivo mais adequado, realizando um pequeno teste de capacidades.

## 3. AVALIAÇÃO DE DESEMPENHO

Após a avaliação da sensibilidade tátil [Benedito10], dez telemóveis diferentes foram categorizados e utiliza-

dos para avaliar o desempenho dos utilizadores. Para isso, foi desenvolvida uma aplicação que indica através da fala e de forma aleatória, qual a tecla que o participante deve pressionar, num dispositivo de 12 teclas. O desempenho de 21 participantes cegos foi avaliado e cada participante interagiu com todos os telemóveis, pressionando duas vezes cada uma das teclas. O tempo entre o estímulo e o pressionar da tecla foram registados (eficiência), com auxílio de um observador. Os erros cometidos pelos utilizadores foram registados por um segundo observador e validados pelas gravações realizadas (eficácia).

Os resultados demonstram que o número de erros e o tempo (Figura 1) variam entre telemóveis. A análise de variância (*one-way ANOVA*) permite-nos afirmar que as características dos teclados afectam significativamente a eficácia ( $F_{9,200}=11941$   $p < .01$ ) e eficiência ( $F_{9,200}=4300$   $p < .01$ ) dos utilizadores.

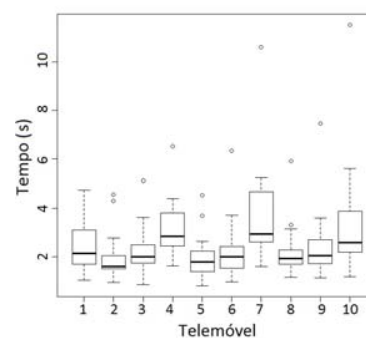


Figura 1: Tempo por telemóvel.

## 4. TRABALHO FUTURO

O próximo passo é relacionar os diferentes níveis de sensibilidade tátil com as características dos teclados dos telemóveis e criar um modelo preditivo representativo do conhecimento adquirido. O principal objectivo será adequar os dispositivos móveis às capacidades dos utilizadores de forma a maximizar o seu desempenho.

## 5. CONCLUSÕES

A sensibilidade tátil é de extrema relevância para os utilizadores cegos. Mesmo soluções móveis para cegos tendem a negligenciá-la. Nós focamo-nos na sensibilidade tátil e em como diferentes níveis tácteis se reflectem em diferentes capacidades de interacção em dispositivos móveis. O objectivo é perceber como as capacidades tácteis se relacionam com as características dos teclados e fornecer uma ferramenta que melhore o desenho e selecção de dispositivos móveis.

## 6. REFERÊNCIAS

- [WHO07] WHO. Magnitude and causes of visual impairment. (2007).
- [Persad07] Persad, U., et al. Characterizing user capabilities to support inclusive design evaluation. *Universal Access in the Information Society*, (2007), 119-135.
- [Benedito10] Benedito, J et al. The Key Role of Touch in Non-Visual Mobile Interaction. *Mobile HCI 2010*.

# Cognimobile: Diferenças Cognitivas e os Dispositivos Móveis

João Oliveira   Tiago Guerreiro   Hugo Nicolau   Daniel Gonçalves

Instituto Superior Técnico / INESC-ID

Rua Alves Redol, 9, 1000-029, Lisboa

jmgdo@ist.utl.pt.com, {tjvg,hman}@vimmi.inesc-id.pt, daniel.goncalves@inesc-id.pt

## Sumário

*Embora dispositivos como os telemóveis assumam um papel cada vez mais importante na vida diária de muitas pessoas, estes continuam a apresentar dificuldades e restrições a populações com necessidades especiais. Os cegos e deficientes visuais em particular, privados de informação visual na qual a maioria dos dispositivos se baseia, necessitam de um esforço cognitivo suplementar na interacção com telemóveis. Apesar de existir interesse em perceber a importância de características humanas na interacção com tecnologia, existe uma grande lacuna no que respeita a estudos que relacionem capacidades cognitivas e o uso de dispositivos móveis por parte de deficientes visuais. Face ao esforço cognitivo superior, na ausência de visão, pretendem-se caracterizar os diferentes tipos de utilizadores de acordo com as suas capacidades cognitivas, de modo a permitir explorar diferentes métodos de interacção e, assim, criar soluções que se adequem ao perfil de cada um.*

## Palavras-chave

*Avaliação, Cognição, Interacção, Dispositivo Móvel, Deficiência Visual*

## 1. INTRODUÇÃO

Os dispositivos móveis, com especial destaque para os telemóveis, fazem cada vez mais parte do quotidiano de cada um de nós. No entanto, e apesar da sua constante evolução, continuam a apresentar grandes dificuldades e restrições a grupos de pessoas com necessidades especiais, como é o caso da população cega [Guerreiro09].

A interacção com telemóveis, assim como qualquer outro dispositivo, exige um esforço cognitivo que, na ausência de um sentido tão fundamental como a visão, se torna muito mais exigente. Um exemplo muito simples deste esforço adicional prende-se na necessidade de memorização da associação entre teclas, e letras ou funções, de um computador ou telemóvel. Uma exigência cognitiva necessária para alguém que não pode simplesmente, e a qualquer momento, confirmar visualmente os botões.

Apesar de existirem estudos que relacionem capacidades cognitivas e o uso de dispositivos, em tarefas como procura e obtenção de informação e introdução de dados [Czaja07], existe uma grande lacuna no que respeita a trabalhos sobre a influência das capacidades cognitivas de um deficiente visual na interacção com dispositivos móveis. A grande maioria dos estudos foca no envelhecimento e não na perda de um sentido vital como a visão.

É nosso objectivo relacionar capacidades de pessoas, a nível cognitivo, com exigências de produtos. Mais precisamente, pretende-se descobrir qual o impacto cognitivo, resultante da interacção de cegos com telemóveis.

## 2. PERCEBER A AVALIAÇÃO COGNITIVA

Os testes que pretendem avaliar a inteligência geral, ou habilidade cognitiva, de uma forma geral avaliam dois

campos distintos: a habilidade verbal e a não-verbal [Aiken04]. Para a avaliação verbal de cegos, as escalas verbais da terceira versão do teste de Wechsler (WISC-III para crianças e WAIS-III para adultos), são geralmente utilizadas. São compostas por subtestes destinados a avaliar o conhecimento verbal, o raciocínio verbal, a memória e o processamento de informação, que podem ter sido adquiridos quer por vias formais (escola) quer por vias informais (família, comunidade) [Nascimento07].

A componente não-verbal pretende avaliar capacidades independentes da língua materna ou cultura, como a habilidade espacial ou velocidade psicomotora [Aiken04]. De entre vários testes destaca-se o Blind Learning Aptitude Test (BLAT), um teste de inteligência não-verbal específico para o uso com deficientes visuais [Aiken04]. Neste teste, baseado nas matrizes progressivas de Raven, o examinado tem como objectivo através do tacto, descobrir uma relação ou encontrar o item que falta de um determinado padrão.

### 2.1 Entrevistas a Peritos

De modo a perceber, num contexto mais prático e da realidade portuguesa, como é feita a avaliação cognitiva de deficientes visuais, foram realizadas entrevistas de carácter semi-estruturado a diferentes psicólogos. Pretendeu-se igualmente interrogar se a falta de um sentido como a visão, obrigará a um esforço cognitivo superior aquando da interacção com dispositivos, tentando compreender em caso afirmativo, quais as características relevantes.

Para este efeito, foram entrevistados dois especialistas, um psicólogo e formador de uma instituição de apoio a

deficientes visuais, e um psicólogo educacional de uma divisão regional de educação. Embora sob perspectivas diferentes, um trabalha mais com jovens enquanto o outro com uma população mais idosa, ambos realizam frequentemente avaliações cognitivas.

Em termos de instrumentos utilizados, as escalas verbais de Wechsler tiveram o maior destaque, sendo considerado um instrumento fiável. Para uma avaliação não-verbal, é utilizado ou o BLAT, ou alguns testes que embora não se encontram padronizados, permitem ter uma ideia do desempenho da pessoa. Relativamente ao pressuposto de que a falta de visão implica um esforço cognitivo adicional na interacção com dispositivos, ambos os peritos foram unânimes em concordar, salientando que deverá existir um esforço suplementar, particularmente a nível da memória de curto prazo e da habilidade espacial.

### 3. COGNIÇÃO E INTERACÇÃO MÓVEL

De forma a atingir os objectivos propostos e relacionar as capacidades cognitivas com a interacção com telemóveis, será utilizada uma metodologia de desenho centrado no utilizador. Com a colaboração de uma fundação de apoio ao cego, diferentes utilizadores serão envolvidos durante o estudo. Serão avaliados e entrevistados, sendo o seu desempenho medido em diversas tarefas.

#### 3.1 Avaliação Cognitiva

De acordo com um estudo efectuado e os depoimentos dos dois peritos entrevistados, a componente verbal dos testes de Wechsler é o instrumento mais utilizado. Assim, este será usado para avaliar capacidades cognitivas como a memória de curto prazo e habilidades verbais.

De modo a complementar a componente verbal, será feita uma avaliação utilizando um instrumento de avaliação não-verbal, o BLAT, referenciado positivamente tanto pela literatura existente como pelos psicólogos. Este teste permite avaliar características como a habilidade espacial e a velocidade de processamento, capacidades que tais como as já referenciadas, se têm mostrado preditivas de desempenho na interacção com dispositivos [Czaja07].

#### 3.2 Avaliação da Tecnologia

Serão feitos testes em diferentes dispositivos móveis com teclados tradicionais. Diferentes estruturas de teclados permitirão sobretudo relacionar o desempenho com a habilidade espacial. O desempenho será medido inicialmente numa tarefa de baixo nível, de aquisição de teclas.

Posteriormente, serão testadas tarefas de alto nível relativamente à introdução de texto e navegação em menus. Tratam-se de duas das principais operações realizadas em telemóveis e onde possivelmente as capacidades cognitivas terão um maior impacto [Guerreiro09].

Serão testados diferentes métodos de introdução de texto, recorrendo a abordagens MultiTapping, onde cada tecla está associada a diversos caracteres, ou de sistemas de navegação [Guerreiro09], onde o utilizador pode percorrer o alfabeto. Adicionalmente irão também ser avaliadas diferentes abordagens de navegação em menus.

### 3.3 Modelo Preditivo

Dado que o uso de métodos empíricos de avaliação incorre em maiores custos, tanto monetários como de tempo, a criação de ferramentas de auxílio ao desenho mostram-se essenciais. Deste modo, irá ser implementada uma ferramenta que mapeie as características cognitivas de uma pessoa, aos métodos de interacção indicados. À semelhança da teoria *Capability-demand* [Persad07], onde requisitos do produto são equacionados com capacidades de utilizador (Figura 1), procura-se criar um modelo preditivo dos resultados obtidos, onde tendo em conta as características de cada indivíduo, serão apresentadas as técnicas de interacção mais aconselhadas e, consequentemente o dispositivo mais acessível.

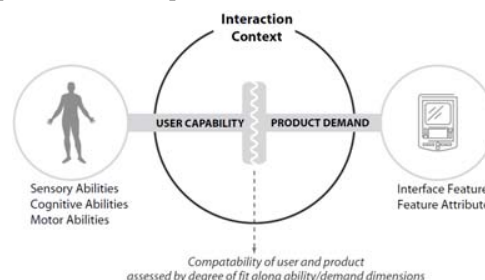


Figura 1: Teoria Capability-demand [Persad07].

### 4. CONCLUSÕES

O uso do telemóvel, cada vez mais vital para a sociedade actual, apresenta um conjunto de desafios a um deficiente visual que implicam um esforço cognitivo superior, podendo restringir ou impossibilitar a utilização destes dispositivos. Diferentes métodos de interacção implicam diferentes exigências cognitivas, pelo que a criação de uma ferramenta de suporte é uma contribuição valiosa para o desenho de interfaces mais inclusivas e adequadas às diferenças individuais de cada utilizador.

### 5. REFERÊNCIAS

- [Aiken04] L. R. Aiken. *Assessment of Intellectual Functioning*. Perspectives on Individual Differences Series. Springer, 2004.
- [Czaja07] S. Czaja e C. Lee. The impact of aging on access to technology. *Universal Access in the Information Society*, 5(4):341-349, 2007.
- [Guerreiro09] T. Guerreiro et. al. NavTap: a long term study with excluded blind users. *Assets '09: Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*, 99-106, New York, NY, USA, 2009. ACM.
- [Nascimento07] E. Nascimento e C. E. Flores-Mendoza. WISC-III e WAIS-III na Avaliação da Inteligência de Cegos. *Psicologia em Estudo*, 12(3):627-633, 2007.
- [Persad07] U. Persad, P. Langdon, e J. Clarkson. Characterising user capabilities to support inclusive design evaluation. *Universal Access in the Information Society*, 6(2):119-135, 2007.

# *ClimbOn* – An infinite game with difficulty adjustment

Ricardo Ameixa  
INSTICC  
Setúbal

rameixa@sapo.pt

Fausto Mourato  
DSI / EST / IPS  
Setúbal

fausto.mourato@estsetubal.ips.pt

João Morais  
ComOn  
Setúbal

joao.morais@comon.pt

## Abstract

*This document describes an infinite game called ClimbOn. The game concept is fairly simple and aims to be suitable for casual gamers that play for short periods, wanting an instant action system with direct and short challenges. The developed system does an automated procedural generation of content with difficult adjustments in order to keep levels challenging to the user and without compromising the final rank system.*

## Keywords

*Casual gaming, Content generation, Human factors*

## 1. INTRODUCTION

As games become more casual and mass oriented, the need to produce new entertaining content quickly enough to satisfy gamers becomes a demanding task for game developers and designers. Procedurally generated levels and difficulty adjustment techniques help companies to achieve these goals, lowering design efforts.

The purpose of this project was the development of a videogame with the application of those techniques in addition to a simple design and game play. The game, called *ClimbOn*, gives to the player the control of an avatar climbing up a building with infinite floors while avoiding falling objects, open windows and other obstacles. The use of procedurally generated content creates a different experience each time the user starts a game. Furthermore, as the user is playing, the game constantly adjusts difficulty to keep the player *on the edge* but not frustrated.

## 2. MOTIVATION

The term casual gaming refers to games that require a low commitment and skill set by the user. This type of games is normally simple in rules and has a low cost of production and distribution. Most of them are playable for free through web browsers and, nowadays, they are also commonly available on mobile phones. Due to the nature of their quick action game play and their vast yet low skilled audience, normally these games don't use difficulty selection mechanics. Such mechanics are commonly associated with more complex ruling sets and higher commitment games. However, a game with a fixed low difficulty setting soon is forgotten by the player, since his/her skills improve as he/she plays it. The addition of dynamic difficulty adjustment (DDA) systems is a good solution to this problem as it allows the game to evolve with the player to a certain degree and, as an added benefit, it also allows more experienced players to enjoy it without feeling excessive easiness.

## 3. RELATED WORK

### 3.1 Procedural generation of gaming content

A popular example of the use of procedural generation and difficulty adjustment techniques is the videogame *Canabalt* [SemiSecret08], developed by Semi Secret Software. This 2D platform game creates segments based on basic buildings and adjusts distances and type of jumps required by the player to how well he/she is performing. The final result is a fluid and immersive experience. A more academic reference worth mention is the work of Compton and Mateas [Compton06], where basic units for 2D platform levels construction are defined as cells. The composition of these cells into patterns and choice of the relations between them define the difficulty of the challenges through the game. By using a physics model and calculating the paths the player can take and based on game play mechanics, the system is able to restrict the levels generated to adequate solutions. The principles presented in this work lead to an implementation presented by Smith et Al. [Smith09].

### 3.2 Difficulty adjustment

Nicollet et al. [Nicollet04] presented an extensive definition of difficulty concepts in dexterity-based games. Among those principles, one can extract that surprise isn't difficulty. This is as a very valued rule set in guiding what variables should and can be adjusted by DDA systems in dexterity-based games. Also, the work of Hunicke and Chapman [Hunicke04] focus the study of DDA and how these systems create immersive experiences. They make use of a probabilistic technique that measures players' performance based on the difficulty he experiences when given an obstacle. Considering Csikszentmihalyi's definition of flow [Csikszentmihalyi90], they start by assessing what states the player should be kept into and they determine the changes that are needed to force a state change without interfering with the player experience.

## 4. TECHNICAL ASPECTS

### 4.1 Game

As previously stated, we developed a game with one single level with one building constituted by infinite floors. The player has to climb it up while avoiding falling objects or other obstacles. Each floor has been represented as a set of three windows, thus this is a configurable parameter. The game automatically increases the player speed upwards, leaving to the user the control of the side to side movement as a way to avoid obstacles, in order to continue playing. The basic score of the player is determined by the distance climbed: the bigger the distance, the bigger the score.

### 4.2 Level Generation

To guarantee game play continuity, the game has to assure that the player has always a valid path to follow. In order to accomplish this goal, the game takes into account that the player is not able to transverse a two window space in just one floor, thus it needs to have its path position either to be placed in the same position or on an adjacent position relatively to the last floor created. In the calculation of the path, we used a scaling progressive percentage factor that can be influenced by the DDA system. A new set of floors is created each time the player enters the current set. Having this buffer gives the game time to analyze the player performance and generate the transition set seamlessly. The obstacles placed outside the path are randomly chosen to create diversity in the facade of the building. Some of the obstacles may even be created as path looking in order to create visual noise as an added difficulty element.

### 4.3 Difficulty adjustments

To be able to adjust difficulty, the game needs a way to measure the player performance. This measurement is done by directly relating closer positioning to the obstacles with the increased difficulty the player has to react to said obstacles. In other words, the game measures the average distance between the player and the obstacles in his neighborhood at all times and in case the player is too close for too long, the DDA system will change the flow of the game. For that, it just needs to change the progressive percentage factor in order to adapt the place where the next path window is more willing to be. As a simple heuristic, keeping the avatar in the same row is easier than moving to an adjacent one. Later adjustments are also done, for instance, adding more or less power ups, changing their power type or increasing the player upwards speed. Having a bigger space between movements on the path or a slower upwards speed, gives the player more time to react.

## 5. RESULTS

In this section we present the result of our game. In figure 1 we can see the main look of the game, which is by this time being tweaked in its design, especially in which concerns to models and textures. It is possible to observe three effective floors generated with low difficulty. With this low difficulty level, only a small amount of obstacles (in this case: windows) were generated.



Figure 1 – Screenshot from our prototype: *ClimbOn*

## 6. CONCLUSIONS AND FUTURE WORK

### 6.1 Conclusions

We presented a game called *ClimbOn* that generates infinite self adjusted levels to provide challenges to casual gamers. This work represents a practical implementation of two main concepts: procedural generation of game spaces and difficulty adjustment. Our primal tests with a restricted set of users showed that the difficulty adjustments are transparent to the user, which realizes that the game gets harder with time but without strong transitions.

### 6.2 Future Work

The main purpose of this article was to present our approach with a concretization in a real game, having the main focus in the procedural generation and the product result. However, we believe that is important to expand some of the principles, in particular to effectively use a metric for difficulty in this game and extract data, such as losing probabilities, among others, from game play in order to extend and validate some of our assumptions.

## 7. REFERENCES

- [Csikszentmihalyi90] Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. NY: Harper Collins, 1991.
- [Compton06] Compton, K., Mateas, M.: Procedural level design for platform games. Proceedings of the 2nd Artificial Intelligence and Interactive Digital Entertainment Conference, 2006.
- [Hunicke04] Hunicke, R., Chapman, V.: AI for dynamic difficulty adjustment in games. Challenges in game artificial intelligence. Pittsburgh: AAAI Press, 2004.
- [Nicollet04] Nicollet, V. Difficulty in Dexterity-Based Platform Games. Gamedev.net. (May 22, 2010). <<http://www.gamedev.net/reference/design/features/platformdiff/>>
- [SemiSecret08] CANABALT: (Browser Game) Semi Secret Software, 2010. Last Accessed: May 22, 2010 <<http://canabalt.com>>
- [Smith09] Smith, G., Treanor, M., Whitehead, J., Mateas, M. 2009. Rhythm-based level generation for 2D platformers. *Proceedings of the 4th international Conference on Foundations of Digital Games* (Orlando, Florida, April 26 -30, 2009). ACM, NY, 175-182.



# Sistema multi-modal de identificação de utilizadores IPTV- um processo de investigação

Telmo Silva  
Departamento de Comunicação e  
Arte, Universidade de Aveiro,  
Portugal  
tsilva@ua.pt

Jorge Trinidad Ferraz de Abreu  
Departamento de Comunicação e  
Arte, Universidade de Aveiro,  
Portugal  
jfa@ua.pt

Osvaldo Rocha Pacheco  
Departamento de Electrónica,  
Telecomunicações e Informática,  
Universidade de Aveiro, Portugal  
orp@ua.pt

**Sumário:** As tecnologias de internet de banda larga de elevada velocidade encontram-se em disseminação crescente no mercado doméstico, assistindo-se, paralelamente, a um contínuo e rápido avanço das tecnologias de distribuição audiovisual. A conjunção destes dois factores contribuiu para o aparecimento do serviço de distribuição de televisão designado por IPTV (Internet Protocol Television). Este conceito de televisão tem associado, em contraponto com a televisão convencional, a bidireccionalidade que permite o desenvolvimento de um conjunto de aplicações interactivas. No entanto, no cenário em que mais do que um habitante partilha a mesma televisão, a automatização desta personalização carece de uma identificação da pessoa/espectador que está em frente ao televisor e não da simples identificação da Set-Top Box existente na habitação.

**Palavras chave**— IPTV, personalização, vídeo, identificação, intrusivo, bidireccionalidade.

## I. INTRODUÇÃO

A televisão é, sem dúvida, uma das invenções que mais mudou as nossas vidas. O seu impacto nas sociedades reflecte-se em diversas vertentes, nomeadamente enquanto agente de socialização e, também, de sociabilização [1] - esta é vista por muitos, como um meio que abriu caminho à construção de opinião e de conhecimento pelas massas e que influencia a nossa forma de encarar a vida em termos de valores, tradições e normas [2].

Fruto da evolução gigantesca que a utilização da Internet tem assumido, torna-se pertinente prestar um olhar atento para os diversos impactos que esta tem vindo a infligir à televisão. Ao nível das audiências televisivas, é certo que o consumo de serviços Web e de fornecimento de vídeo na Internet têm marcado o seu terreno [3], embora diversas componentes, tais como o tamanho de ecrã, a qualidade de imagem e, mesmo, o conforto do sofá, vão mantendo a televisão fortemente enraizada na 'dieta' mediática dos portugueses (Figura 1) [4].

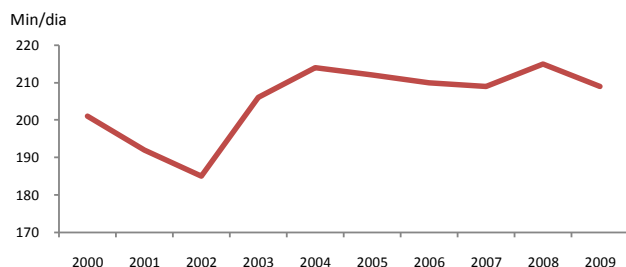


Figura 1- Tempo de visionamento de televisão de 2000 a 2009 [4]

Do ponto de vista do posicionamento do mercado televisivo, verifica-se que a televisão, à semelhança de muitos outros meios de comunicação, deslocou parte da sua influência e

presença para a Internet com a implementação de portais (com informação complementar sobre as emissões, concursos, etc.). No que toca à componente infra-estrutural, a Internet também é, cada vez mais, utilizada como suporte à difusão de conteúdos. É neste contexto tecnológico que surge a IPTV (Internet Protocol Television) [5], um serviço de Televisão digital fornecido através de uma rede IP, a mesma que suporta a Internet. Este conceito de difusão de conteúdos, difere do sistema de TV tradicional pois implica que os espectadores tenham uma ligação dedicada a um operador de telecomunicações. Neste contexto, e porque esta ligação garante bidireccionalidade de transmissão de informação, a oferta de conteúdos e serviços interactivos é significativamente facilitada[6]. No entanto, no cenário actual em Portugal, e na maioria dos casos, a identificação de quem está a interagir com a televisão e que, consequentemente, pode beneficiar de elevados índices de personalização de serviços, peca por esta ser efectuada ao nível da casa (através dos identificadores únicos das respectivas Set-Top boxes) e não do indivíduo que, realmente, está em frente ao televisor. Esta forma de 'identificar utilizadores' é, assim, pouco flexível e limitada, pelo que se justifica o desenvolvimento de um sistema que facilite e automatize a identificação do(s) utilizador(es) que, na realidade, esta(ão) em frente ao televisor. É neste quadro que se insere o interesse deste estudo - plataforma multi-modal, não intrusiva, de identificação de utilizadores de IPTV. A concretização destes mecanismos de identificação de utilizadores, permitirá otimizar a experiência de utilização através de serviços personalizados e adequados ao contexto [7]. Aqui inserem-se, por exemplo, experiências como: i) escolha automática do canal mediante aprendizagem prévia dos hábitos do utilizador; ii) serviços de alerta para redes de cuidados de pessoas idosas; iii) serviços de tele-

assistência a pessoas idosas; iv) monitorização de hábitos de pessoas com mobilidade reduzida.

## II. QUESTÃO DE INVESTIGAÇÃO

A metodologia apresentada neste artigo baseia-se na preconizada por Quivy e Campenhoudt [8]. Neste sentido, a primeira fase da investigação consistiu em delinear uma questão que norteie todo o desenvolvimento do trabalho. Assim, para perceber como a questão da identificação de utilizadores IPTV é encarada pelos respectivos utilizadores, bem como a sua resposta pode ser determinante no desenvolvimento de novos serviços, este trabalho de investigação contempla a especificação, desenvolvimento e teste de um protótipo que facilite o estudo da seguinte questão de investigação: “Quais as características do binómio techno-social que permitem sustentar um processo não intrusivo de identificação de utilizadores IPTV?”.

## III. OBJECTIVOS

Tendo este trabalho como objectivo geral “perceber qual o melhor processo de identificação (não intrusiva) de utilizadores IPTV para fomentar o desenvolvimento de serviços personalizados de televisão interactiva”, importa definir um conjunto de objectivos mais específicos que permitam demarcar, de uma forma mais concreta, o trabalho a realizar. Assim, os objectivos específicos do trabalho são:

- Fundamentar a pertinência do desenvolvimento do protótipo através de um estudo exploratório que caracterize a população portuguesa face à aceitação e utilização de serviços de IPTV, para perceber: i) qual a percepção em relação às vantagens de ter conteúdos e serviços personalizados; ii) qual a percepção sobre a utilização de um sistema de identificação automática;
- Conceptualizar, desenvolver e testar um protótipo funcional de um sistema de identificação de utilizadores IPTV, com o intuito de analisar: i) o comportamento dos utilizadores perante um sistema de identificação; ii) o tipo de dados que deve ser armazenado sobre cada utilizador para garantir a sua identificação; iii) as vantagens de fornecer conteúdos personalizados, como por exemplo serviços noticiosos, sugestão de canais, serviços de governo e cidadania contextualizados, entre outros; iv) novos serviços baseados no sistema de identificação; a percepção dos utilizadores sobre o grau de intrusão do sistema.

## IV. OPERACIONALIZAÇÃO DA INVESTIGAÇÃO

No trabalho aqui apresentado, a questão de investigação sugere múltiplas abordagens na procura de uma solução para o problema da identificação dos utilizadores de serviços IPTV. Numa fase embrionária da investigação, e depois de uma revisão da literatura e do levantamento do estado da arte, pretende-se efectuar um conjunto de entrevistas exploratórias para desenhar, ainda que de forma empírica e algo incipiente, o protótipo a desenvolver [8]. Estas entrevistas permitirão perceber qual a percepção que os utilizadores têm do grau de intrusão de um sistema de identificação. As entrevistas exploratórias, em conjunto com a leitura de artigos científicos, ajudarão a construir a problemática de investigação. Enquanto

os artigos ajudam a perceber e situar o problema de partida, as entrevistas facilitam a percepção dos aspectos a ter em conta e que podem alargar, ou rectificar, o campo de investigação [8]. Para desenvolver o protótipo serão comparadas múltiplas tecnologias como *bluetooth*, RFID (Radio-Frequency IDentification), redes Wi-Fi, reconhecimento de imagem, tanto ao nível da fiabilidade, facilidade e flexibilidade de utilização como ao nível de ingerência na privacidade do telespectador. A Figura 2 ilustra, de uma forma sucinta e esquematizada, o processo a seguir.

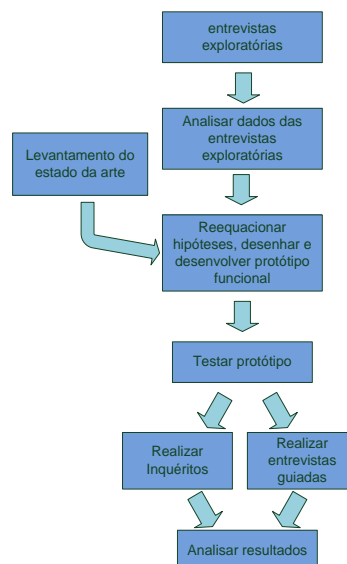


Figura 2 - Operacionalização da metodologia de investigação

A operacionalização da investigação aqui apresentada implica múltiplas fases, sendo que, nas fases em que é necessário recolher e analisar dados, os métodos a utilizar serão a observação directa, entrevista semi-dirigida e o inquérito por questionário [8].

## V. REFERÊNCIAS

1. Wolton, D., *Penser la communication*. 1997, Paris: Flammarion.
2. Abreu, J. and V. Branco, *Os rumos e as faces da Televisão Interactiva*, in *Futuro da Internet*, C. Atlântico, Editor. 1998.
3. Ribeiro, Â.A., *A televisão e a concorrência digital: o fim do monopólio do vídeo*. Estudos em Jornalismo e Mídia, 2007.
4. Obercom (2009) *Perspectivas de Implementação da Televisão Digital em Portugal Caracterização do Acesso TV 2008*
5. Cooper, W. and G. Lovelace, *IPTV guide - delivering audio and video over broadband*. 2006, informtv.
6. Cesar, P., K. Chorianopoulos, and J.F. Jensen, *Social television and user interaction*. Comput. Entertain., 2008. 6(1): p. 1-10.
7. Neuwirth, O., et al., *Towards a New User Experience in IPTV: Convergence Services and Simpler E-commerce on IMS-based IPTV*. NEC Technical Journal 2008. 4.
8. Quivy, R. and L.V. Campenhoudt, *Manual de Investigação em Ciências Sociais*. 2005: Gradiva.

# Support for inferring user abilities for multimodal applications

Carlos Duarte Pedro Feiteira Daniel Costa David Costa

LaSIGE and Informatics Department

Faculty of Sciences of the University of Lisbon

Campo Grande, 1749-016 Lisboa, Portugal

{cad,pfeitoira}@di.fc.ul.pt, {thewisher,dcosta}@lasige.di.fc.ul.pt

---

## Abstract

*Multimodal interaction offers users the possibility to interact with computational systems using natural interaction modes. However, different users interact differently, integrating modalities in different patterns. By knowing how users exploit different modalities, we can try to build adaptive systems that will increase the effectiveness of multimodal interaction. For that we need to characterize user abilities. However, this is a time and effort consuming process, particularly in the early stages of design. We propose an approach that can be used in the early stages of the design process, by resorting to a Wizard of Oz based prototype of a multimodal system.*

## Keywords

*Multimodal interaction, Adaptive interfaces, Adaptive multimodal systems, Design, Prototyping, Wizard of Oz.*

---

## 1 INTRODUCTION

Multimodal interaction, by offering alternative input and output modalities, provides users with more natural interaction choices. By resorting to modalities like speech and gestures, they bring interaction closer to what users are accustomed from human-human interaction [Dumas 09]. Additionally, users can benefit from the flexibility of having different alternatives to interact. If the user can't point to an object to select it (physical impairment) speech recognition can replace the selection modality employed. Thus, natural interaction modes and the flexibility to allow users to select the most appropriate modes for them, are two of the main advantages of multimodal interfaces.

However, in order to take full advantage of this flexibility, and the associated variability in the context of use it brings, we need to assist the multimodal application. It is easy to envision various factors, sensory, physical and cognitive, that can impact the use of the different available modalities. Besides these, other factors, related to social or environmental characteristics, can also have such an impact. In order to tackle the added complexity of managing this number of variables, we can endow the system with capabilities to adapt to its user, thus becoming a multimodal adaptive system. As a result, not only users can choose the modalities they use to interact, but also the system can select the more appropriate modalities to present information to the users, and even improve the performance of recognition based modalities for input [Oviatt 08].

In order to support adaptation, the system will need to be able to identify its users and build a user model for each one. This requires a system that is capable of identifying

the user's abilities, both for its use and for the construction of the user model. To allow a multimodal adaptive system to have access to this information and to use it effectively from the beginning, some kind of initialization process is required. However, during the early stages of design, when there is not a clear definition of the user model attributes yet, to develop such an initialization procedure, using multimodal interaction, would require an effort disproportionate to the results that could be attained from it.

To support achieving the goal of characterizing the user, without developing a full scale multimodal application, we propose a *Wizard of Oz* based approach, allowing a human operator to replace some of the recognizers, while still supporting the goal of assisting in understanding what user characteristics are relevant for the application in development and how different users combine different modalities.

## 2 CONTEXT

Although the aforementioned motivation and concerns are universal, this approach is being developed in the scope of *GUIDE*, an EU funded project. *GUIDE* deals with multimodal interaction tailored for the elderly population in a home setting, with a TV as the central hub. However, the interaction means to be made available to the user are not limited to a TV screen, sound system and remote control.

Different modalities will be considered in *GUIDE*'s scope. Input will be provided through the remote control, speech, gestures, with and without the support of a surface, and recognition on a tablet based computing device. Output will be presented on the TV screen, and also through sound, speech and haptic feedback. Additionally, video and audio output will be combined in the form of avatars

to increase the personalization effect and try to create an emotional attachment between users and system.

### 3 THE PROTOTYPE

The *Wizard of Oz* based prototype aims at collecting enough information to build the initial version of the user profile. Although the profiling application will be used only once by each user, it has the responsibility of being the first contact users will have with the whole system, and thus contribute to their initial impressions of it. To acquire the required information, a set of different tests will be performed, involving several input and output modalities. The prototype's results will support two types of conclusions: first, how users behave in the face of what is requested from them; second, if the proposed tests are necessary and sufficient to collect all the required information.

#### 3.1 User Abilities Characterizing Tests

As stated before, the aim of the application is to characterize the user's abilities. To work towards this aim, the prototype will have to allow the evaluation of tests in three main areas: sensory, physical and cognitive. The sensory tests will focus on visual and auditory related senses. To evaluate the user's vision impairments tests targeting visual acuity, and any impairments relating to text contrast and color [Ishihara 17] will be ministered.

Physical tests will enable the characterization of the user's physical abilities. In order to be able to support the most natural interaction possible, a full characterization of the user's abilities regarding gestures is the aim. To do that, we need to characterize the gesture making abilities regarding different attributes and different body parts. Tests will be made to evaluate the capability of making gestures with the head, the upper body and the lower body. The attainable gesture's amplitude, height and rotation will be evaluated.

The final set of tests aims at characterizing the cognitive traits that will be important when interacting with the set of applications considered in the scope of *GUIDE*. These traits have been initially identified as memory, working memory and recognition speed, and the set of tests used will be based on the Cambridge university's CANTAB tests [Sahakian 92].

#### 3.2 Modalities of the Prototype

To be able to execute and gather data from all the tests described, the application will have to rely on video and speech recognition for input. As argued before, developing such an application to validate early stages of design would impose high effort demands. Our approach is to employ a *Wizard of Oz* based prototype that will replace the recognition based modalities with a human *Wizard*.

In what concerns the output modalities, text, video, speech and sound cover the requirements from the tests. These outputs are all controlled by the *Wizard* through the prototype's interface. The *Wizard* can also control the flow of the application's execution (what test to administer next), as well as the presentation parameters. The interface also

includes a set of fields to allow the *Wizard* to report and comment on the tests' results.

One feature of the prototype is the use of an avatar to interact with the user. The responsibility of asking the user to perform tasks and give feedback about the user's actions is placed on the avatar and not on a human experimenter. We opted for incorporating the avatar since the early stages of prototyping, since it is important to understand if, by using a humanized figure to interact, we are able to strengthen the engagement of elderly users with such a system.

#### 3.3 Automated vs. Manual Tasks

As stated above, the interaction with the application is controlled by the *Wizard*. However, for some of the tests it is extremely hard for the *Wizard* to accurately perceive the user's input. This is particularly true for pointing tasks. It is virtually impossible for a human to accurately perceive the point that is being pointed at on a TV screen by other person. Our solution was to integrate a custom made application that uses a *WiiMote* as the input device, and allows for presenting on the screen a cursor showing what is being pointed at. The cursor can appear only on the *Wizard's* screen or on both the *Wizard's* and the user's screens.

### 4 CONCLUSION

In this paper we presented our approach for designing an application for the profiling of users of a multimodal adaptive system. Multimodal systems are dependent on different recognition technologies to support natural interaction modes. Given that, both the individual recognition technologies and their integration are still expensive, in terms of man-power to develop, and prone to errors in individual recognition modalities, our approach rests on a *Wizard of Oz* based prototype. Using a *Wizard* to replace the recognition technologies we eliminate possible recognition errors while postponing the integration of the multiple recognition technologies to a latter stage in the design and development of the system. For the user profiling task it is not relevant that the recognition is technologically based, neither that the recognition technologies are integrated.

### References

- [Dumas 09] B. Dumas, D. Lalanne, e S. Oviatt. Multimodal interfaces: A survey of principles, models and frameworks. Em *Human Machine Interaction*, páginas 3–26. Springer Berlin / Heidelberg, 2009.
- [Ishihara 17] Shinobu Ishihara. *Tests for colour-blindness*. Hongo Harukicho, Tokyo, 1917.
- [Oviatt 08] S. Oviatt, C. Swindells, e A.M. Arthur. Implicit user-adaptive system engagement in speech and pen interfaces. Em *CHI*, páginas 969–978, 2008.
- [Sahakian 92] B.J. Sahakian e A.M. Owen. Computerized assessment in neuropsychiatry using cantab: discussion paper. *J. of the Royal Society of Medicine*, 85:399–402, 1992.

# Avaliação Pericial de Barreiras ao Acesso sobre Sítios Web de Entidades Públicas

Nádia Fernandes  
Universidade de Lisboa  
nadia.fernandes@di.fc.ul.pt

Rui Lopes  
Universidade de Lisboa  
rlopes@di.fc.ul.pt

Luís Carriço  
Universidade de Lisboa  
lmc@di.fc.ul.pt

## Sumário

*A acessibilidade de sítios Web é um factor crucial para pessoas com deficiências, para que estes consigam aceder a informação relevante existente em páginas Web. Este artigo apresenta uma análise pericial efectuada ao sítio Web do Governo de Portugal (<http://www.portugal.gov.pt>) baseada na metodologia de detecção de barreiras ao acesso Barrier Walkthrough. Mostramos que este tipo de metodologias potencia a detecção de um número maior de problemas, comparativamente aos processos normalmente utilizados.*

## Palavras-chave

*Acessibilidade Web, Avaliação Pericial, Barrier Walkthrough.*

## 1. INTRODUÇÃO

O conceito de acessibilidade baseia-se na facilidade de acesso a conteúdos ou serviços por parte de pessoas com algum tipo de incapacidade sem que para isso necessitem do auxílio de terceiros.

A directiva em Portugal que obriga à acessibilidade de sítios Web institucionais é a Resolução do Conselho de Ministros número 155/2007 [PCM07], que utiliza a norma WCAG 1.0 [Chisholm99]. De acordo com esta directiva, “a organização e apresentação da informação facultada na Internet pelos sites do sector público”, devem ser escolhidas para permitir ou facilitar o seu acesso pelos cidadãos com necessidades especiais. A acessibilidade deverá abranger, no mínimo, a informação relevante para a compreensão dos conteúdos e para a sua pesquisa”.

Apresentamos um estudo da acessibilidade no sítio do Governo de Portugal, para utilizadores: invisuais, daltónicos e com deficiências nos membros superiores.

## 2. TRABALHO RELACIONADO

### 2.1 Avaliação Pericial da Acessibilidade Web

A avaliação pericial da acessibilidade da Web pode ser realizada tendo em conta que é definido em normas como a WCAG, ou seguindo metodologias de análise pericial como a Barrier Walkthrough [Brajnik09].

A Barrier Walkthrough é uma metodologia em que estão definidas categorias de utilizadores e as barreiras para cada categoria. Sendo as barreiras qualquer condição que impeça um utilizador com uma deficiência de cumprir um objectivo. Assim, para cada categoria de utilizadores são verificadas as barreiras presentes que o impeçam o seu acesso ao sítio. Podendo verificar-se se os critérios da WCAG 2.0 estão a ser cumpridos.

### 2.2 Avaliações Institucionais

A Resolução do Conselho de Ministros número 155/2007 indica que se deve assegurar que a informação disponibilizada pela Administração Pública na Internet seja acessível a cidadãos com necessidades especiais. Na actualidade, o acesso às tecnologias da informação e da comunicação e a capacidade para a sua utilização são diferenciadores das oportunidades sociais.

O logótipo de "Certified Accessibility" da UMIC [ASC09] indica a acessibilidade do sítio em que se encontra. É dinâmico e permite a vigilância dos conteúdos de um sítio na Web, através do validador eXaminator (utiliza WCAG 1.0). Apresenta vários estados de acordo com o cumprimento das normas. Esta avaliação é automática, não detectando muitos dos erros detectados com o Barrier Walkthrough, nem os agrupando por deficiências. A existência de mais de 600 mil pessoas com incapacidades em Portugal (EU 2002) que são “info-excluídas”, suportam a ideia de que criar sítios Web acessíveis.

Em 2008, foram realizados estudos de acessibilidade dos sítios Web das mil maiores empresas portuguesas em volume de negócio [Gonçalves09] (INE 2007), seguindo-se as normas do W3C, pelo Grupo de Negócio Electrónico. Os resultados foram: 9,4% apresentam o nível A, uma tem o nível de AA e nenhuma apresenta o nível máximo.

## 3. AVALIAÇÃO PERICIAL

### 3.1 Metodologia

Seguiu-se a metodologia Barrier Walkthrough, para as categorias de utilizadores/deficiências já referidas. Escolheram-se vinte e cinco templates representativos dos conteúdos e estruturas do sítio do governo. O critério de escolha dos templates escolhidos foi a diferença da estrutura entre eles e a verificação visual de barreiras.

## 4. RESULTADOS

A avaliação resultou na detecção de diversos problemas de acessibilidade. Obteve-se uma taxa de aprovação de 30% para invisuais, de 50% para daltónicos e de 50% para deficiências dos membros superiores.

De seguida vão ser descritas algumas das barreiras encontradas e possíveis formas de eliminação das mesmas.

### 4.1 Links genéricos

São links que não fornecem informação suficiente para que se compreenda o seu conteúdo. Por exemplo, um link com o texto “mais...” que nos redirecciona para uma página de notícias (Figura 1). Seleccionar-se-ia este problema modificando os labels de links para que dessem pistas da página que vai ser aberta.

```
<a href="/pt/GC18/Noticias/Pages/20100608_Not_CM_PMEInveste.aspx"> mais...</a>
```

Figura 1 – Exemplo no código de um link genérico

### 4.2 Objectos opacos

São componentes totalmente opacos para os leitores de ecrã. Neste caso, são utilizados vídeos em Flash (Figura 2). Isto poderia ser resolvido garantido que o objecto é acessível, seguindo directivas específicas, caso isso não fosse possível o objecto deveria ser removido.

```
swfobject2.embedSWF(
"/pt/GC18/ConteudosTransversais/Flashes/20091118_
Governo/slideshowpro.swf",
"flashcontent", "978", "255", "9.0.0", false,
flashvars, params, attributes);
```

Figura 2 – Exemplo no código de um objecto opaco

### 4.3 Eventos do rato

São eventos desencadeados apenas com a utilização do rato, pessoas que não utilizassem o rato nunca conseguiriam utilizá-los. Isto verifica-se na utilização de event handlers ("onclick", ...) que são orientados para o rato. O problema era resolvido utilizando event handlers lógicos ("onfocus", ...) além dos orientados para o rato.

### 4.4 Nova janela

Quando se clica num link somos redireccionados para uma nova janela no browser, devido à utilização de 'target="\_blank"'. O problema seria resolvido evitando abrir-se novas janelas. Se for mesmo necessário deve haver um link ou botão que permita fechar a janela, para que os utilizadores percebam que se abriu uma nova janela e que têm a possibilidade de a fechar.

### 4.5 Não é possível saltar links

Não é permitido saltar directamente para o conteúdo da página. Por exemplo o utilizador tem de passar por todos os links anteriores antes de chegar ao link que pretende.

### 4.6 Contraste visual insuficiente

A página contém elementos cujo contraste entre estes e o fundo é insuficiente. Por exemplo, o contraste entre fundo branco e ícones/texto cinzento. Este problema poderia ser resolvido com o aumento do contraste dos elementos.

## 5. DISCUSSÃO

Apesar dos esforços para tornar o sítio do governo mais acessível, encontraram-se várias barreiras de acesso, que impedem os utilizadores de aceder correctamente aos conteúdos e de realizarem operações. Por exemplo, os invisuais poderiam ter de percorrer toda a página para encontrar um link e pessoas com deficiências nos membros superiores não evitariam a utilização do rato.

Sendo as metodologias empregues insuficientes, deveriam ser tomadas mais medidas para que o site fosse acessível a todos os seus utilizadores. A resolução conselho ministros, não está a ser cumprida e está desactualizada, utilizando ainda WCAG 1.0, quando a norma suportada pelo do W3C é a WCAG 2.0.

## 6. CONCLUSÕES E TRABALHO FUTURO

A acessibilidade na Web é importante para que as pessoas com deficiências consigam compreender conteúdos e realizar as actividades que pretendem. Com a análise pericial realizada ao sítio do Governo de Portugal, baseada na metodologia Barrier Walkthrough, verificou-se que este tem bastantes barreiras ao acesso.

No seguimento deste trabalho, aplicaremos este processo no estudo de outros sítios Web de instituições públicas. Procederemos à comparação quantitativa de processos de avaliação pericial com avaliadores automáticos e à exploração destas limitações no âmbito de testes de acessibilidade com utilizadores finais.

## 7. REFERÊNCIAS

- [ASC09] Agência Para A Sociedade Do Conhecimento, Programa Acesso, 20 de Setembro de 2009, [http://www.acesso.unic.pt/webax/nota\\_tecnica\\_logo.html](http://www.acesso.unic.pt/webax/nota_tecnica_logo.html)
- [Brajnik09] Brajnik, Giorgi, Barrier Walkthrough Heuristic Evaluation Guided by Accessibility Barriers, Março de 2009, <http://users.dimi.uniud.it/~giorgio.brajnik/projects/bw/bw.html>
- [Chisholm99] Chisholm, Wendy, Vanderheiden, Madison Gregg, Jacobs, Ian, Web Content Accessibility Guidelines (WCAG) 2.0, 5 de Maio de 1999, <http://www.w3.org/TR/WCAG10/>
- [Gonçalves09] Gonçalves, Ramiro, Pereira, Jorge, Martins, José, Mamede, Henrique, Santos, Vítor, Web Ponto de Situação das Maiores Empresas Portuguesas, Setembro 2009, [http://www.acesso.unic.pt/estudos/1000maioresempresas\\_apdsi\\_0909.pdf](http://www.acesso.unic.pt/estudos/1000maioresempresas_apdsi_0909.pdf)
- [PCM07] Presidência Do Conselho De Ministros, Resolução Do Conselho De Ministros N.º 155/2007, Diário Da República, 1.ª série — N.º 190 — 2 de Outubro de 2007, <http://www.unic.pt/images/stories/publicacoes200710/RCM%20155%202007.pdf>



# NavTilt: Interface gestual para cegos

David Lucas    Hugo Nicolau    Tiago Guerreiro    Joaquim Jorge

Instituto Superior Técnico / INESC-ID

Rua Alves Redol, 9, 1000-029 Lisboa

david.lucas@ist.utl.pt, {hman, tjvg}@vimm.inesc-id.pt, jaj@inesc.pt

## Sumário

*A crescente miniaturização dos dispositivos móveis e as suas interfaces visualmente exigentes impõem diversos desafios à população cega. Em particular, os métodos de introdução de texto tradicionais mostram-se desadequados às necessidades destes utilizadores. Este artigo descreve uma nova abordagem de entrada de dados em dispositivos móveis com base numa interface gestual. O NavTilt apresenta-se como um método de interacção simples e natural, recorrendo à utilização de apenas uma mão, podendo ser usado sem retorno visual.*

## Palavras-chave

*Dispositivo Móvel, Introdução Texto, Gestos, Cego, Telemóveis.*

## 1. INTRODUÇÃO

Devido à evolução da tecnologia, os dispositivos móveis têm vindo a ficar cada vez mais pequenos, com maiores capacidades de processamento e armazenamento. Aliado ao seu baixo custo, estes dispositivos tornaram-se ubíquos, sempre presentes, possibilitando a sua utilização em qualquer local. Segundo a ITU (*International Telecommunications Union*), no início de 2009 existiam 4,6 mil milhões de subscrições de telemóveis em todo o mundo, o que corresponde a 67% da população [Bud-09].

Estes dispositivos têm um papel cada vez mais importante nas sociedades modernas, quer a nível pessoal ou social, muitas pessoas não conseguem imaginar a sua vida sem eles. Apesar desta grande evolução e difusão de aparelhos móveis, alguns problemas ainda subsistem. Em particular, muitos utilizadores cegos vêem-se impedidos de usar os seus dispositivos em situações do dia-a-dia. As soluções tradicionais, como os leitores de ecrã, encaram a deficiência visual sob uma perspectiva limitada. Esta solução apenas oferece aos utilizadores o conteúdo do ecrã sob a forma de retorno auditivo. No entanto existem outros problemas.

Na tarefa de introdução de texto, por exemplo, o utilizador não tem qualquer retorno da distribuição das letras pelas teclas, necessitando de um maior esforço cognitivo. Por outro lado, devido à ausência de retorno visual, os utilizadores são obrigados a confiar nas suas capacidades tácteis para conseguir “sentir” e seleccionar as teclas pretendidas. Apesar do estereótipo do utilizador cego, com capacidades sobrenaturais, esta não é a realidade. De facto, 82% da população cega tem mais de 50 anos, resultando numa diminuição exponencial do tacto [WHO07].

Para além das suas capacidades tácteis os utilizadores têm de fazer uso de ambas as mãos para conseguir manipular o seu dispositivo. Tipicamente, a mão não dominante segura-o enquanto a mão dominante executa a

aquisição das teclas. Este tipo de interacção torna-se desadequado a diferentes situações diárias, em especial no caso da população cega, que usualmente transporta uma bengala, cão guia ou outros objectos.

Face aos problemas que as interfaces móveis actuais possuem, relativamente à introdução de texto, foi desenvolvido um novo método que proporciona uma baixa exigência motora (é usada apenas uma mão), táctil e visual.

## 2. A REALIDADE ACTUAL

Com vista a perceber as dificuldades que a população cega sente em relação à introdução de texto em dispositivos móveis, foi efectuada uma análise de tarefas. Esta análise foi dividida em três componentes: questionário inicial, observação e entrevista semi-estruturada.

Com base num questionário realizado a 20 utilizadores de um centro de formação para deficientes visuais (Fundação Raquel e Martin Sain), foi verificado que todos eles possuem telemóveis tradicionais, sendo que 90% usa leitor de ecrã. Pôde-se verificar também que 80% dessas pessoas usam o serviço de mensagens escritas, sendo um método de comunicação habitual.

A fase de observação foi realizada a 5 utilizadores com o objectivo de perceber as suas dificuldades nas tarefas de entrada de dados. Foi-lhes pedido que escrevessem algumas frases no seu telemóvel, usando o método tradicional (MultiTap) e leitor de ecrã. As frases consistiam em ditados populares, de forma a minimizar a necessidade de memorização. Caso os utilizadores fizessem uso de ambas as mãos neste tipo de tarefas, era-lhes pedido que a repetissem usando apenas uma mão.

Todos os participantes necessitaram de ambas as mãos para escrever texto nos seus dispositivos. Verificou-se ainda que quando lhes era pedido para usarem apenas uma das mãos, estes sentiam muitas dificuldades. De facto, todos eles necessitaram de usar algum suporte para pousar o dispositivo. Houve mesmo dois participantes

que se recusaram a executar a tarefa por acharem que não eram capazes.

Após a observação foi efectuada uma entrevista semi-estruturada com todos os participantes, onde estes foram encorajados a referir as dificuldades sentidas. Para além da interacção com apenas uma mão, os participantes referiram que por vezes têm dificuldades em identificar as teclas dos seus dispositivos (devido ao baixo relevo) e a elevada exigência cognitiva do método tradicional (memorização da distribuição das letras pelo teclado)

### 3. NAVTILT

Com base nos problemas identificados foi criado o método *NavTilt*. Este diferencia-se dos métodos de introdução de texto tradicionais, baseados em abordagens *multitapping*, por possuir uma interacção baseada em gestos 3D e uma nova organização do alfabeto.

Os gestos são uma parte importante do *NavTilt*. Estes oferecem uma forma de interacção natural, com baixa carga visual e podem ser executados com apenas uma mão. No entanto, a associação de cada letra a um gesto iria tornar o sistema complexo e traria dificuldades de aprendizagem para o utilizador. Por essa razão foi utilizada também uma representação do alfabeto sob a forma de matriz [Guerreiro08]. Este modelo (Figura 1) já demonstrou as suas vantagens de aprendizagem e utilização quando comparado com o método tradicional.

Com esta representação, o número de erros é drasticamente reduzido, visto que o utilizador tem apenas de saber o alfabeto. Não existem caminhos errados, apenas caminhos mais curtos para cada letra. Utilizadores mais inexperientes podem começar na letra 'A' e ir navegando horizontalmente até chegar à letra pretendida. Por outro lado, utilizadores mais proficientes podem começar a aprender novos caminhos e usar atalhos, navegando verticalmente entre as vogais.

Através da junção deste modelo com a utilização de gestos simples de navegação, criou-se uma sinergia em que é possível introduzir texto usando apenas uma mão, com baixa carga cognitiva e visual, não sendo necessária a localização constante de teclas.

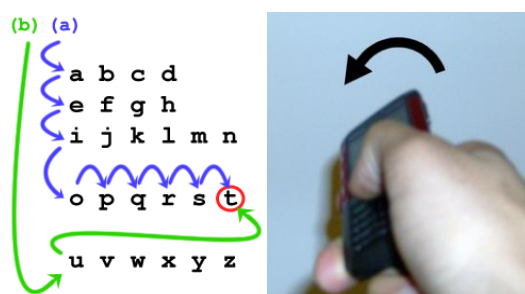


Figura 1. Matriz de navegação (esquerda); Gesto de rotação (direita)

Analisando os gestos de uma forma mais pormenorizada, pode-se observar que estes têm como função uma navegação pelo alfabeto em quatro direcções. Assim sendo, foi escolhida uma abordagem semelhante à usada em [Wigdor03] para desambiguação de caracteres: inclinação

ou rotação do dispositivo. Mais precisamente o inclinar para a esquerda ou direita, faz navegar para a esquerda ou direita na matriz, respectivamente, e o inclinar para a frente ou para trás, faz navegar para baixo ou para cima, respectivamente. Um exemplo de inclinação poderá ser visto na Figura 1.

Para iniciar a interacção gestual o utilizador terá de pressionar uma tecla de activação e inclinar o dispositivo na direcção pretendida. Para aceitar a escolha da letra basta apenas largar a tecla. Sempre que o utilizador escolher um carácter irá ouvi-lo.

Caso o utilizador queira aceitar uma palavra basta pressionar a tecla de activação sem navegar em qualquer direcção. Automaticamente, será lida a palavra e introduzido um espaço em branco. Para além da tecla de activação gestual, o botão esquerdo das teclas direccionais permite eliminar o último carácter. Sempre que é apagada uma letra esta é sintetizada.

No que diz respeito ao equipamento necessário, este método apresenta-se como uma solução de baixo custo, sendo apenas necessário um acelerómetro.

### 4. CONCLUSÕES E TRABALHO FUTURO

Neste artigo apresentamos algumas das principais dificuldades sentidas por utilizadores cegos em tarefas de introdução de texto em dispositivos móveis. Nomeadamente, a elevada exigência motora, tátil e cognitiva do método tradicional.

Perante as dificuldades observadas propomos um novo método, *NavTilt*, que usa gestos tridimensionais, proporcionando uma interface natural e simples que pode ser usada apenas com uma mão e sem recurso ao retorno visual.

Como trabalho futuro, o *NavTilt* irá ser avaliado e comparado com outras técnicas e modalidades de interacção. Pretendemos validar o uso de interfaces gestuais para a população cega e observar o processo de aprendizagem destes utilizadores quando comparado com o método tradicional.

Adicionalmente, o sistema será ainda avaliado com utilizadores sem deficiências em situações de mobilidade. Com estes testes pretendemos verificar se utilizadores normo-visuais poderão beneficiar deste método, em situações visualmente exigentes, onde têm de usar a sua atenção para monitorizar o ambiente envolvente (e.g. navegar numa plataforma de metro). Em caso afirmativo estaremos mais próximos de criar interfaces mais inclusivas ou unificadas que poderão ser usadas por ambas as populações.

### 5. REFERÊNCIAS

- [Budde09] Budde, P. Global Mobile Communications – Statistics, Trends and Forecasts. Tech. rep., BuddeComm, 2009.
- [Guerreiro08] Guerreiro, T. et al. NavTap and Braille-Tap: Non-Visual Texting Interfaces. RESNA'08, Arlington, VA, 2008.
- [Wigdor03] Wigdor, D. Balakrishnan, R TiltText: using tilt for text input to mobile phones, 2003.

# Email Visualization: A Context-Based Approach

Celso Cardoso  
ccrc26@gmail.com

Paula Alexandra Silva  
paula.silva@fraunhofer.pt

Fraunhofer AICOS  
Rua do Campo Alegre 1021  
4169-007 Porto, Portugal

## Abstract

*Email has become a widespread communication medium, but its interface has not undergone many changes over the years. The main contribution of this work is to propose a context-based email visualization application, which goal is help users acquire habits of email management and search.*

## Keywords

*Information Visualization, Human-Computer Interaction*

## 1. INTRODUCTION

Nowadays, most people communicate using one or more email accounts, leading to an increasing amount of sent and received emails. This fact, combined with email persistency (i.e. email messages may remain on accounts for long periods of time), increases the difficulty to find particular information.

Over the past years, multiple features (such as instant messaging) have been added to email clients such as Microsoft Outlook, Mozilla Thunderbird, Gmail, among others. However, none of these changes addressed the user interface paradigm. There are several email visualizations (eg.: Anymails [Horn07] and Mail Garden [Wilkens08]), their purpose is not the same as that of an email system, as these only allow a limited set of actions to be performed.

The main goals of this work were the analysis of existing email visualizations, a brief study of the user's needs (for our study, college students), the development of a new visualization application to help users search email messages and to make them acquire email organization habits by avoiding the accumulation of email messages in the inbox. Low and medium-fidelity prototypes were developed and evaluated. The visualization developed presents a different design while also incorporating the main functionalities of an email client. A different feature is the fact that the email client is context-based.

## 2. RELATED WORK

Most email clients explore the information of emails regarding the relations between contacts, threads and keywords, among other criteria. Recently, new visualization concepts have emerged, providing new metaphors for email systems. Anymails and Mail Garden are examples of such metaphors. Anymails, created by Carolin Horn in 2007 [Horn07], represents received emails as colored microbes. The microbes "swim" freely across the screen with different speeds, according to their status (read, un-

read or replied). Mail Garden, proposed by Ken Wilkens in 2008 [Wilkens08], represents emails as trees in a forest. Each tree represents an email and its height the email size. On both visualizations, the presentation of information related to a certain email is triggered by hovering the mouse over an entity representing it.

## 3. METHODOLOGY

This work followed a user-centered design approach [Abras04]. Data collection was accomplished through semi-structured interviews [Unger09]. This technique helped us to understand users' habits and difficulties while using their usual email system. This type of interviews enabled the authors to deepen some interesting points of the conversation. After collecting the information from the interviews, a low-fidelity prototype [Rudd96] was developed. After the prototype conclusion, several users were subjected to a usability test [Dumas99]. With the results collected from this test, a redesign was made to improve the identified problems. When the redesign was finished, another usability test was made with different subjects, allowing the authors to evaluate other functionalities that were not previously tested and to verify if the improvements were correctly implemented. During the test, the subjects were asked to "report out loud" or "say out loud" any thoughts, feelings and opinions experienced during the task [Senger93]. Also, during the test, we simulate the real behavior of the application. In the end of each usability test, the participants were asked to complete a questionnaire in order to collect information about the importance of some functionalities and the usability of the application. Questionnaires were also performed at the beginning of each of the interviews and usability tests in order to gather demographic data.

## 4. USER RESEARCH AND DATA COLLECTION

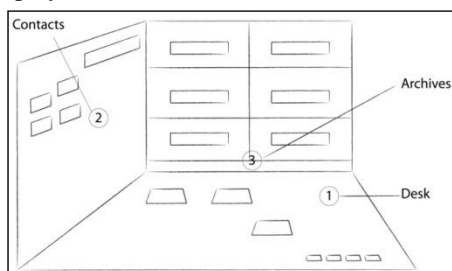
College students are the target audience of the developed visualization approach, since the largest uptick in email use occurs when students enter college and then this ef-

fect continues as these students graduate and enters the workforce [Stewart09].

For this study, seven students attending their undergraduate or masters degrees studies were interviewed. To compare results, four professors that receive a large number of emails were also interviewed, since these people experience more problems while managing their emails. Results showed that the differences between the two groups that stood out the most were the ability to organize email into folders and the management of the received emails by the professors. The students did not exhibited the habit to organize emails into folders, had difficulties to find a specific email and did not relate too much information to their contacts. Usually, the students only related the name of the contact. The main problem of the professors is the large amount of received emails causing them to have difficulties organizing and reading all their email messages.

## 5. PROTOTYPE AND EVALUATION

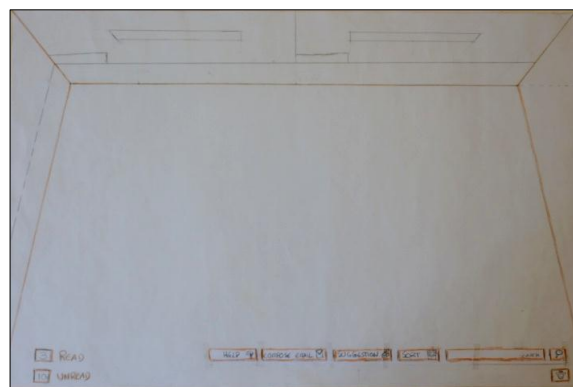
Figure 1 shows a sketch of the email visualization approach that we developed and that uses the visualization metaphor of an office desk. The application is composed by one table (1 in Figure 1), where received email messages are disposed, and two walls. One wall is composed by archives (3 in Figure 1), which contain folders, where the messages can be organized. The other wall (2 in Figure 1) displays the contacts.



**Figure 1 - Email visualization sketch**

Functionalities, considered to be beyond the normal email functionalities, are suggestions (that pop-up whenever a subject is repeated over a number of emails) and the different areas to view contacts, folders or incoming messages. The letters are sorted from the left upper corner to the bottom right corner by received date. The user can also sort the letters and create piles by sender or date. He can also drag and arrange them at will. When the user has too many letters on the table, these start accumulating on a pile. When the number of archives exceeds the visible area, a scroll bar appears. The user can search for the archive name by using the search option on the contacts wall menu. Through this menu the user can also delete or create archives or folders. To access archives and folders the user just has to double-click. The user can move letters between folders and archives dragging them or through the option "Move". It is possible visualize the letters by its sender. The user can also go to the contacts wall menu and visualize the received letters from a given contact. The paper prototype shown in Figure 2 was tested with twelve persons, divided in two usability tests. In the first test, the main problems were related to deleting

messages, moving messages between folders and identifying some actions that were more hidden. The application was redesigned in order to solve these problems and explore other aspects. The second test showed that the changes made had been correctly implemented.



**Figure 2 - Paper prototype**

## 6. DISCUSSION AND CONCLUSIONS

Interviews provided helpful information to the prototyping phase. The feedback obtained from the usability tests was important, considering the subject's opinions and problems identified. Low-fidelity and medium fidelity prototypes allowed quick modifications.

## 7. REFERENCES

- [Abrás04] C. Abrás, D. Maloney-Krichmar, and J. Preece, "Encyclopedia of Human-Computer Interaction," in *User-Centered Design [Draft]*.: Thousands Oaks: Sage Publications, 2004.
- [Dumas99] J. Dumas and J. Redish, "Introducing Usability Testing," in *A Practical Guide to Usability Testing*.: Intellect Ltd; Rev Sub edition, 1999, pp. 22-38.
- [Horn07] C. Horn, "Natural Metaphors for Information Visualization". Boston, Massachusetts College of Art in Boston, Master of Arts in Design, 2007.
- [Rudd96] J. Rudd, K. Stern, and S. Isensee, "Low vs. high-fidelity prototyping debate," *Interactions*, pp. 76-85, 1996.
- [Senger93] C. Senger, "Thinking Aloud Protocols: A Diagnostic Tool for Teaching Legal Problem Solving," in *Thomas M. Cooley Law Review*., 1993, pp. 367-382.
- [Stewart09] M. Stewart. (2009) MediaPost. [Online]. [http://www.mediapost.com/publications/?fa=Articles.showArticle&art\\_aid=115417](http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=115417)
- [Unger09] R. Unger and C. Chandler, "User Research," in *Project Guide to UX Design, A: For user experience designers in the field or in the making*.: New Riders, 2009.
- [Wilkins08] Kjen Wilkins. (2008) Mail Garden. [Online]. <http://human-centered-visualizations.com/>

# Time Automaton

## A visual mechanism for temporal querying

Luís Certo  
certo@fe.up.pt

Teresa G. Dias  
tgalvao@fe.up.pt

José Borges  
jlborges@fe.up.pt

Department of Industrial Engineering and Management  
School of Engineering of the University of Porto, Rua Dr. Roberto Frias, s/n 4200-465, Porto

### Abstract

*Visual querying mechanisms are normally comfortable and accessible mechanisms. However, they generally do not provide the means for formulating a large set of different types of temporal interrogations. In this paper we propose the Time Automaton, a highly flexible visual language, capable of formulating many types of temporal queries. Time Automaton's logic is innovative and unlike available visual mechanisms it is not based on the direct conversion of visual queries into a predefined set of commands of a textual querying language, like for example, SQL.*

### Keywords

*temporal query, visual mechanism, visual model, regular expression*

## 1. INTRODUCTION

Using visual querying mechanisms is normally more comfortable than using traditional text-based querying languages. Most important, this type of mechanisms is useful when users do not possess programming skills. However, querying temporal data is a complex process, and, to the best of our knowledge, available visual querying mechanisms [Jin09] [Shneiderman92] [Edsall97] do not provide the means for formulating a significant set of different types of temporal interrogations.

Available visual mechanisms convert visual configurations directly into commands of traditional text-based languages [Snodgrass84] [Snodgrass95]. Since the operators of textual querying languages were not designed to be visually manipulated their adaptation to the logic of visual mechanisms is not natural and it generates functional limitations.

In this paper we propose the Time Automaton, a highly flexible visual mechanism that is capable of formulating a vast set of different temporal queries. Time Automaton's logic is quite different from the logic of available visual mechanisms, and unlike them, Time Automaton does not use any available text-based querying language to execute its queries. Time Automaton principles are simple and the algorithm that executes the Time Automaton queries is simple to implement. A specific data structure is required to work with the mechanism. Such structure is uncomplicated and an easily implementable algorithm can automate its creation.

## 2. DATA STRUCTURE

Time Automaton requires temporal data to be encoded as a string. We call this data structure the **temporal string**. Two types of words compose the temporal string, **anchors** and **facts**. Anchors define the beginning of temporal moments and facts encode temporal data records, that is, the data itself. The temporal string uses a *prefix* notation where each anchor affects the anchors and facts that are subsequently

positioned.

For the artificial dataset represented in Table 1 the corresponding temporal string, with a month granularity and considering the year season for each month, can be defined as follows:

year,2009 month,Jan,Winter **facts,a,b**  
month,Feb,Winter **facts** month,Mar,Spring  
**facts,c,d,e,f** month,Apr,Spring **facts**  
month,May,Spring **facts** month,Jun,Summer  
**facts** month,Aug,Summer **facts,g,h**  
month,Sep,Summer **facts** month,Oct,Autumn,  
**facts** month,Nov,Autumn **facts,k,l,m,n**  
month,Dec,Autumn **facts**  
year,2010 month,Jan,Winter **facts,p,q**  
month,Feb,Winter **facts** month,Mar,Spring  
**facts,s**

DATE	DATA_RECORDS
Jan, 2009	a,b
Mar, 2009	c,d,e,f
Aug, 2009	g,h
Nov, 2009	k,l,m,n
Jan, 2010	p,q
Mar, 2010	s

**Table 1: Artificial dataset containing temporal information**

## 3. MECHANISM DESCRIPTION

A Time Automaton query is represented as an oriented graphs with the following properties: **1** - One of the nodes is the root, which has no incident edges. **2** - Every other node is defined by an expression, which corresponds to a word in the temporal string. **3**. Edges can have a value to define the maximum number of traversals.

Each node's expression is in fact a regular expression [Aho80]. A regular expression is written in a formal language that is interpreted by a regular expression engine, which analyzes a text and finds which parts match the provided specification. For example, a regular expression



engine can analyze a text in attempt to find a word with a specific syntax, and starting the search at a specified index.

In order to abstract the process of writing the regular expressions that define nodes, some predicates were created: **1. Fact()** corresponds to `'facts((?:\w+)+)?'`. It defines the syntax of a fact and specifies that it must be captured [Friedl06]. That is, facts are stored in a **capture buffer**, where they can be accessed separately from other matched words. **2. Undef()** is the same as `\w+`. It is used to match any word, indifferently. In the temporal querying context this predicate is useful for creating queries that do not specify characteristics of temporal moments. For example, in the query *"all facts occurred in every month of Summer"* the months' names are not specified. **3. Neg(charact)** corresponds to `'(?:!bcharactb)\w+')`. In the temporal querying context this predicate is useful for creating queries that exclude certain temporal moments like *"all facts occurred in every month except March"*.

The algorithm that executes Time Automaton queries corresponds to a simple graph traversal algorithm. Every time a node is visited its regular expression is supplied to a regular expression engine that finds the next matching word in the temporal string. All matched words that are stored in the **capture buffer** correspond to the facts that answer the temporal query. As an example, for the temporal string in Section 2 and the question *"which facts occurred in March 2009?"*, the corresponding Time Automaton query can be formulated as represented in Figure 1.

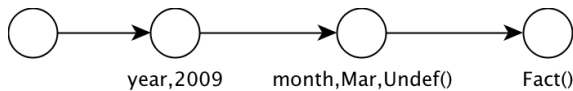


Figure 1 - Example of a Time Automaton query

Time Automaton allows nodes to have multiple outgoing edges. In such case, only one edge is traversed, and the selection criteria is to traverse the edge linking to the node that corresponds to nearest word. In terms of performance, searching for every word for each possible next node is not a good approach. A better approach is to construct a regular expression joining all regular expressions of all possible next nodes, i.e., if two regular expressions  $r1$  and  $r2$  are joined they form a regular expression  $'r1 | r2'$ , where  $'|'$  is the boolean operator OR. Then, the regular expression engine searches the next word that matches any of the elementary regular expressions. After finding this word the algorithm knows which node should be visited next. As an example of a multiple outgoing edge situation, for the temporal string in Section 2 and for the question *"which facts occurred in 2009 in all months except March"*, the corresponding Time Automaton query is represented as depicted in Figure 2.

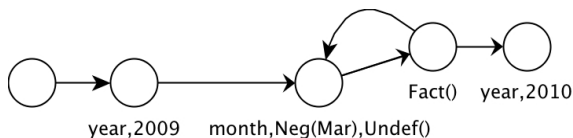


Figure 2 - Example of a Time Automaton query including nodes with multiple outgoing edges

Since Time Automaton visual layout is based on graphs it is simple to create a graphical interface that implements the mechanism. Moreover, the predicates created to encapsulate regular expressions can be easily mapped into interface

controls, enabling regular expressions to be graphically built. As an example, Figure 3 shows a screenshot of an interface we've created to work with the Time Automaton.

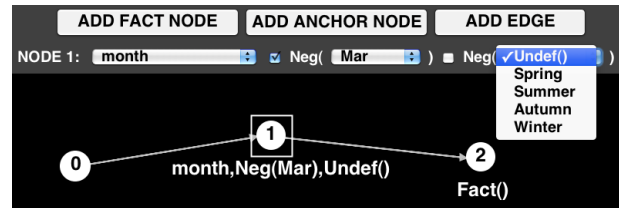


Figure 3 - Interface that implements the Time Automaton

#### 4. CONCLUSIONS

In this paper we present the Time Automaton, a highly flexible visual mechanism that is capable of formulating many types of temporal queries. As a matter of fact, Time Automaton enables the formulation of a type of queries, which are ordinal queries (e.g. *"all facts occurred in every second Saturday of every month"*), which in SQL [Codd83] are not directly formulable.

A Time Automaton query is represented as an oriented graph in which nodes are defined with regular expressions. In order to abstract the process of writing such expressions some predicates were devised. This abstraction layer and the fact Time Automaton's layout is based on graphs makes the mechanism simple to implement in a graphical interface.

Time Automaton's logic is innovative and, unlike available visual mechanisms, Time Automaton does use available text-based querying language to execute its queries. The algorithm that executes the queries is a graph traversal that follows some simple rules, thus, it is straightforward to implement. Since Time Automaton queries are graphs it is simple to store and reuse them.

#### 5. REFERENCES

- [Aho80] A. Aho. Pattern matching in strings. Formal Language Theory: Perspectives and Open Problems, pp. 325-347, 1980.
- [Codd83] E. F. Codd. A relational model of data for large shared data banks. Commun. ACM, 26(1), pp. 64-69, 1983.
- [Friedl06] J. Friedl. Mastering regular expressions. O'Reilly Media, Inc., 2006.
- [Jin09] J. Jin and P. Szekely. Querymarvel: A visual query language for temporal patterns using comic strips. Visual Languages and Human-Centric Computing, pp. 207-214, 2009.
- [Edsall97] R. Edsall, D. Peuquet, A graphical user interface for the integration of time into GIS, Proceedings of the 1997 American Congress of Surveying and Mapping Annual Convention and Exhibition, Seattle, WA, 1997, pp. 182-189.
- [Shneiderman92] C. Ahlberg, C. Williamson, and B. Shneiderman. Dynamic queries for information exploration: An implementation and evaluation. Conference on Human Factors in Computing Systems, Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 619 - 626 , 1992.
- [Snodgrass84] R. Snodgrass. The temporal query language tquel. In PODS '84: Proceedings of the 3rd ACM SIGACT-SIGMOD symposium on Principles of database systems, pp. 204-213, New York, NY, USA, 1984. ACM.
- [Snodgrass95] R. T. Snodgrass, editor. The TSQL2 Temporal Query Language. Kluwer, 1995.



# Um protótipo de um sistema de apoio à decisão em actividades logísticas de localização e distribuição

Rui Borges Lopes  
CIO/DEGEI Univ. de Aveiro  
Aveiro  
rui.borges@ua.pt

Carlos Ferreira  
CIO/DEGEI Univ. de Aveiro  
Aveiro  
carlosf@ua.pt

Beatriz Sousa Santos  
IEETA/DETI Univ. de Aveiro  
Aveiro  
bss@ua.pt

## Sumário

*Duas das mais importantes actividades logísticas são a determinação da localização de equipamentos ou serviços e a gestão das rotas de distribuição. Uma correcta gestão destas actividades é um factor crítico para o sucesso de empresas e instituições. Este trabalho pretende mostrar um protótipo de um sistema de apoio à decisão (SAD) desenvolvido para estas actividades logísticas (profundamente relacionadas).*

*O protótipo permite a obtenção de dados, interacção com web map services (WMS), resolução de instâncias e visualização de resultados. O principal objectivo é a obtenção de um conjunto de boas soluções para estes tipos de problemas; no entanto, subjacente ao seu desenvolvimento está também a concretização de uma interface de utilizador intuitiva e eficaz, permitindo a sua utilização por uma audiência mais alargada.*

## Palavras-chave

*Apoio à decisão, localização, distribuição, localização-distribuição, web map services.*

## 1. INTRODUÇÃO

A localização de equipamentos e distribuição de produtos são duas das actividades logísticas mais importantes. Estas actividades têm vindo a ser estudadas de uma forma quantitativa na área da Investigação Operacional (IO). Os problemas matemáticos inerentes são designados de problemas de localização (PL) [Hale10] e problemas de rotas de veículos (PRV) [Díaz10]. Ainda que as decisões inerentes se refiram a horizontes temporais distintos (decisões referentes a localização são tipicamente estratégicas, e por isso de longo prazo, enquanto que decisões de distribuição são tipicamente efectuadas numa base diária, logo de curto prazo) estão intimamente relacionadas, podendo/devendo por isso ser analisadas de forma integrada. Em IO o problema integrado foi apelidado de problema de localização-distribuição (PLD) [Nagy07].

Neste trabalho pretende-se mostrar um protótipo de um SAD que permita que utilizadores sem conhecimentos específicos em IO, ou sobre os modelos matemáticos inerentes, sejam capazes de obter soluções para estes problemas (PL, PRV e PLD), que são de difícil resolução (NP-difícil).

Igualmente utilizadores com conhecimentos específicos na temática podem ter vantagens na utilização do protótipo, visto que permite: inserir dados de forma fácil; a obtenção de dados reais,

devido à interligação com WMS; e visualizar graficamente os resultados (possibilitando perceber mais facilmente como os diferentes algoritmos funcionam, ajudando na correcta definição de parâmetros).

O protótipo desenvolvido tem a possibilidade de incorporar algoritmos avançados ao mesmo tempo que permite que o decisor possa de uma maneira fácil, rápida e intuitiva obter e comparar um conjunto de boas soluções.

## 2. DESCRIÇÃO DA APLICAÇÃO

A aplicação desenvolvida concentra, num ecrã principal, as principais funcionalidades, as quais se acedem através da barra de ferramentas ou com recurso ao menu.

O motivo desta escolha prende-se com a necessidade de uma interface de fácil utilização e aprendizagem por parte dos utilizadores alvo desta aplicação. Considera-se que o utilizador alvo utilizará a aplicação esporadicamente sendo tipicamente um decisor com: formação superior; moderados conhecimentos informáticos e uma forte experiência na temática (experiência profissional em situações reais de instalação de depósitos, definição de rotas de distribuição e/ou desenho de sistemas logísticos).

Este perfil do utilizador alvo sugere que a ênfase do projecto da interacção com o utilizador deve ser sobretudo na facilidade de aprendizagem.

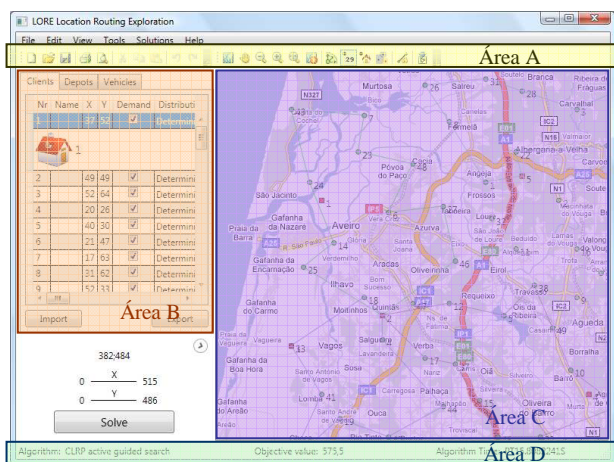
Com este objectivo e tendo em consideração as tarefas que pretendem realizar com a ajuda do protótipo, assim como os principais princípios e directivas que se referem às questões fundamentais do projecto de interfaces de utilizador (como consistência, compatibilidade, familiaridade, *feedback*, robustez, etc.) [Mayhew92, Dix93], foi desenvolvida a interface com o utilizador.

O protótipo foi desenvolvido para plataformas Windows, usando uma metodologia orientada a objectos (*Unified Modelling Language* – UML), e implementado em XAML (*Windows Presentation Foundation*), sendo a linguagem *code-behind* o C#.

## 2.1 Interface

A interface, ao longo do seu desenvolvimento, tem sido sujeita a avaliações com utilizadores, com vista a aferir a sua usabilidade. A interface é constituída por quatro partes distintas (Figura 1):

- uma barra de ferramentas que permite o acesso a diversas funcionalidades (Figura 1, Área A);
- uma grelha de dados onde consta a informação referente aos problemas (Figura 1, Área B);
- uma área de visualização onde é possível visualizar (e editar) informação relativa aos mapas (Figura 1, Área C);
- uma barra de estado com informação relativa ao algoritmo utilizado e solução obtida (Figura 1, Área D).



**Figura 1: Interface de utilizador caracterizada pelas diferentes áreas de visualização.**

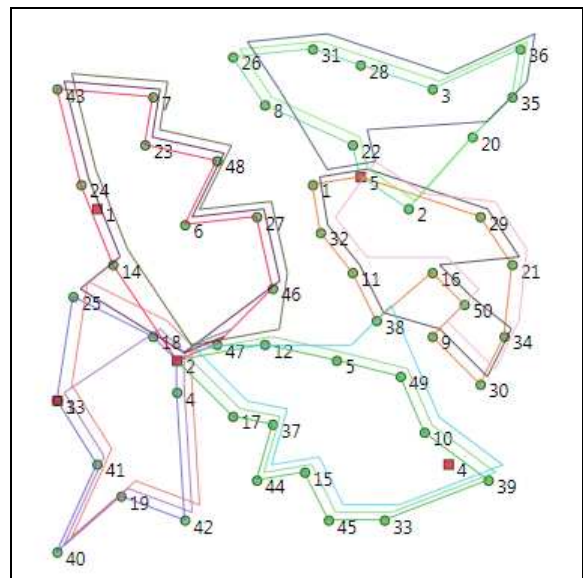
## 2.2 Algumas Funcionalidades

Actualmente o protótipo permite:

- introduzir (ou editar) dados novos (ou já existentes) por forma a caracterizar um problema;
- interagir com WMS, por forma a obter informação geográfica em tempo real;
- obter soluções e visualizá-las de forma numérica ou gráfica;

- comparar visualmente diferentes soluções;
- interagir com o utilizador no processo de obtenção de soluções (capitalizando a sua experiência e prevendo eventuais restrições não contempladas no modelo);
- guardar (exportar) os dados criados para ficheiros facilmente interpretáveis (XML).

A título de exemplo, uma das funcionalidades desenvolvidas, é a possibilidade de comparar graficamente diferentes soluções (Figura 2). Assim, o utilizador pode facilmente comparar soluções, permitindo reduzir a dimensão do problema (e respectiva complexidade) se encontrar rotas comuns a todas as “boas” soluções.



**Figura 2: Comparação entre três soluções.**

Como trabalho futuro aponta-se a integração com sistemas de informação geográfica (SIG) e a incorporação de novos algoritmos na aplicação.

## 3. REFERÊNCIAS

- [Díaz10] Díaz, B.D. The VRP Web. Networking and Emerging Optimization, University of Malaga, 2010. <http://neo.lcc.uma.es/radi-aeb/WebVRP/>
- [Dix93] Dix, A., Finlay, J., Abowd, G. e Beale, R. *Human-Computer Interaction*. Prentice Hall, 1993.
- [Hale10] Hale, T. Trevor Hale's Location Science References. University of Houston, 2010. <http://gator.dt.uh.edu/~halet/>
- [Mayhew92] Mayhew, D.J. *Principles and Guidelines in Software User Interface Design*. Prentice Hall, 1992.
- [Nagy07] Nagy, G. e Salhi, S. Location-routing: issues, models and methods. *European Journal of Operational Research*, 177 (2007), 649–672.

# m-Tourism 2.0: os sistemas integrados de multiagentes para apoio à informação e comunicação no turismo

Pedro Beça  
Dep. Comunicação e Arte, UA  
Aveiro  
pedrobeca@ua.pt

Rui Raposo  
Dep. Comunicação e Arte, UA  
Aveiro  
raposo@ua.pt

## Sumário

*A convergência tecnológica, nas mais variadas áreas e actividades, tem contribuído para o despertar de áreas de investigação e desenvolvimento num número de esferas científicas e económicas cada vez mais abrangentes e transdisciplinares. Neste contexto de convergência, impossível de dissociar de conceitos como a inovação, a criatividade, competitividade, sustentabilidade e participação, o turismo tem registado uma multiplicidade de mutações na sua relação com as Tecnologias de Informação e Comunicação reconhecendo as mesmas como um valor acrescentado a compreender e a integrar. Esta relação com a tecnologia, cuja área de enquadramento tem sido gradualmente reconhecida internacionalmente como a do e-Tourism, tem vindo a confrontar-se com novos desafios, como o m-tourism, resultantes da emergência dos dispositivos móveis. O presente artigo tem como objecto a descrição de um trabalho de investigação, presentemente em curso, no qual são estudados os sistemas integrados de multiagentes para apoio à informação e comunicação no m-Tourism e no qual se apresentará o conceito de m-Tourism 2.0. São apresentados, neste âmbito, o enquadramento do estudo, os seus objectivos e algumas considerações finais sobre os potenciais spin offs esperados do projecto.*

## Palavras-chave

*e-Tourism, m-Tourism, WEB 2.0, m-Tourism 2.0, cultura participativa.*

## 1. INTRODUÇÃO

O turismo, nos últimos anos, tem tido um crescimento contínuo, tornando-se actualmente a maior indústria do mundo [Yueh07]. A World Tourism Organization prevê que, até ao ano de 2020, as viagens de turismo internacional tenham um aumento próximo dos 200% [UN World Tourism Organization08]. O turismo tornou-se numa área de negócio altamente competitiva a nível mundial e, tal como noutros sectores de actividade, um dos factores de vantagem competitiva está, cada vez mais, fortemente relacionado com a utilização das Tecnologias de Informação e Comunicação (TIC) neste sector.

## 2. E-TOURISM, M-TOURISM E TOURISM 2.0

É neste contexto de influência das TIC no sector do turismo que se enquadra o e-Tourism, um conceito abordado pela primeira vez no âmbito de uma iniciativa da UNCTAD's [United Nations Conference on Trade and Development02]. *task force on sustainable tourism for development* que tem como objectivo fomentar o desenvolvimento e o uso das TIC com o intuito de fortalecer e desenvolver as capacidades dos países promoverem os seus produtos turísticos.

No âmbito do e-Tourism, tem-se recorrido ao uso de tecnologias cada vez mais inovadoras e mais integradas nos hábitos diários dos turistas. Neste esforço de integração, os dispositivos móveis, tais como, PDAs e *smartphones*, estão a tornar-se instrumentos cada vez mais comuns no dia-a-dia das pessoas e a assumir especial relevância dadas as suas características de portabilidade, mobilidade e autonomia [Pinheiro08], bem como a sua personalização [Huijnen06]. É neste contexto que surge o conceito de *mobile tourism*, também referido como m-Tourism [Huijnen06], um desafio relativamente novo na área do e-Tourism [Kimber05] que tem vindo a projectar-se em diversas vertentes, tais como, no apoio a deslocações e no acesso a informação contextualizada com a localização. No contexto actual, surge ainda um outro fenómeno, a Web 2.0, que é pertinente analisar uma vez que está fortemente relacionado com a questão dos dispositivos móveis e da personalização. Também no Turismo, a Web 2.0 tem vindo a fornecer mecanismos que permitem que os turistas participem livremente nas actividades de criação de produtos turísticos e personalizem esses mesmos produtos. Tem-se, assim, vindo a verificar uma mudança na forma como as entidades relacionadas com o turismo se apresentam aos

turistas, uma vez que muitas destas também passaram a conceber produtos “centrados no consumidor” [Hao08] e com a participação dos consumidores. Para se categorizar esta nova tipologia de produtos turísticos surgiu o conceito de Tourism 2.0.

### 3. DEFINIÇÃO DO PROBLEMA

A pesquisa realizada pelo European Media Laboratory, no âmbito do projecto “*Creation of User-friendly Mobile services Personalized for Tourism (CRUMPET)*”, [revelou que 78% dos utilizadores no mundo valoriza a utilização de serviços móveis no turismo e 60% estão dispostos a pagar pela sua utilização. Verifica-se, no entanto, que muitos desses serviços não estão ainda orientados para o turista. Verifica-se ainda que existem condicionantes relacionadas com: i) o acesso aos conteúdos pretendidos, por parte dos utilizadores; ii) um custo de comunicações reduzido; iii) a grande diversidade de dispositivos móveis, implicando que, para que os conteúdos sejam visualizados correctamente em todos os dispositivos, seja necessária a adaptação dos conteúdos a todos dispositivos; iv) o facto de os serviços de m-Tourism terem diferentes especificações consoante os diferentes cenários de utilização (ar livre, museus, etc.); v) e um insuficiente número de estratégias capazes de ajudar a fomentar a criação de comunidades de utilizadores participativos em torno de serviços de m-Tourism. O trabalho de investigação aqui descrito surge, assim, na sequência de se considerar que é necessário estender a visão sobre o m-Tourism, nomeadamente contribuir para uma visão mais profunda sobre as condicionantes referidas anteriormente sobre a temática do m-Tourism e sobre a sua relação com a Web 2.0 e cultura participativa. Considerou-se, assim, que seria útil a definição de um novo conceito, m-Tourism 2.0 (Figura 1), que representasse a ideia de integração de serviços Web 2.0 e da cultura participativa, em sistemas móveis de apoio ao turismo.

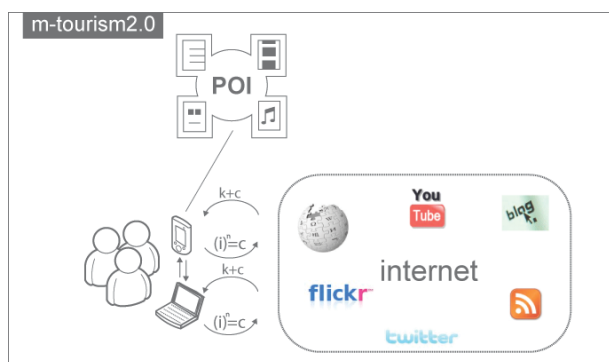


Figura 1 – m-Tourism 2.0

### 4. PRINCIPAIS RESULTADOS E CONTRIBUTOS

O estudo apresentado pretende contribuir para:

- o estudo das implicações de um sistema de m-Tourism 2.0 nos processos de informação e comunicação associados a uma experiência turística;
- o estudo das estratégias a utilizar na criação e actualização dos conteúdos multimédia disponibilizados num sistema de m-Tourism;

- a promoção do turista como um membro da comunidade *prosumer* disposto a enriquecer e a partilhar a sua experiência.

### 5. CONSIDERAÇÕES FINAIS

O trabalho, embora ainda em curso, detém desde já um potencial que se considera interessante, dado que a área do m-Tourism ainda se encontra numa fase de exploração na procura de contextos de uso, serviços e mais-valias que traz para o turismo. Espera-se que o trabalho descrito possa, nesse âmbito, contribuir para a investigação em questões ligadas: à mobilidade do utilizador num espaço e numa experiência de cariz turístico; ao acesso a informação *in context* e *on the spot*; à personalização e partilha da experiência; à disponibilidade de rede; à facilidade de utilização; e à interacção social. Acredita-se que, numa lógica correlacionada com a tendência actual da procura da ubiquidade tecnológica em diversos contextos, a área do m-Tourism 2.0 poderá encontrar um nicho de mercado com potencial na exploração das características do perfil *digital native* cada vez mais presente nos turistas.

### 6. REFERÊNCIAS

- [Hao08] Hao, Z., Wei, Z., & Wenjing, W. (2008). The Mass Customization of Tourism Products Based on WEB2.0: A Collaboration Model by Both Enterprises and Tourists. *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference*, 1-5.
- [Huijnen06] Huijnen, C. (2006). Mobile tourism and mobile government - an inventory of European projects. Retrieved Technical Report, from <http://www.mgovworld.org/whitepapersandcase studies/invenory-mobile-government.pdf>
- [Kimber05] Kimber, J., Georgievski, M., & Sharda, N. (2005). Developing Usability Testing Systems and Procedures for Mobile Tourism Services. *Annual Conference on Information Technology in the Hospitality Industry, HITA 2005*, 12-23.
- [Pinheiro08] Pinheiro, S. C. V., Meiguins, B. S., Meiguins, A. S. G., & Almeida, L. H. (2008). A Tourism Information Analysis Tool for Mobile Devices. *Proceedings of the 2008 12th International Conference Information Visualisation*, 264-269.
- [UN World Tourism Organization08] UN World Tourism Organization. (2008). Tourism Highlights. *Tourism Highlights 2008*. 2008, from [http://www.world-tourism.org/facts/eng/pdf/highlights/UNWTO\\_Highlights08\\_en\\_HR.pdf](http://www.world-tourism.org/facts/eng/pdf/highlights/UNWTO_Highlights08_en_HR.pdf)
- [United Nations Conference on Trade and Development02] United Nations Conference on Trade and Development. (2002). ICT and tourism: a natural partnership. Retrieved 01/05/2009, 2009, from <http://www.unctad.org/Templates/Page.asp?intIdemID=360>

## **Índice de Autores**

Aveiro, 13, 14 e 15 de Outubro de 2010



# Autores

- Alda Lopes Gançarski, 89  
 Alfredo Ferreira, 21, 29, 255  
 Ana Margarida Almeida, 283  
 Ana Paula Cláudio, 159  
 Ana Pedroso, 281  
 Ana Veloso, 263  
 André F. S. Barbosa, 181  
 André Luís, 263  
 Aurélio Nogueira, 281
- Beatriz Sousa Santos, 47, 73, 137, 307  
 Bruno de Araújo, 21, 129  
 Bruno Miguel Cardoso, 193
- Carlos Duarte, 13, 59, 297  
 Carlos E. Silva 81  
 Carlos Ferreira 73, 137, 307  
 Catarina Mendonça, 141  
 Celso Cardoso, 303  
 Cláudia Mourato Nunes, 229
- Daniel Costa 297  
 Daniel Gonçalves, 147, 155, 163, 221, 241, 289, 291  
 Daniel Mendes, 255  
 Daniela da Cruz, 89  
 Daniela Fonte, 89  
 David Costa, 297  
 David Lucas, 301  
 Diogo Mariano, 21  
 Duarte Teixeira, 65
- Emília Duarte, 113
- Fausto José Mourato, 173, 293  
 Fernando Moreira da Silva, 229  
 Filipe Gil, 159  
 Francisco Providência, 7  
 Francisco Rebelo, 113, 229  
 Frutuoso G. M. Silva, 181, 189
- Gabriel Barata, 155, 163  
 Gelly Rodrigues, 281  
 Guilherme Campos, 141  
 Gustavo Zurita, 245
- Hugo Nicolau, 221, 237, 241, 289, 291, 301
- Isabel Sá, 245, 249
- Jennifer Preece, 5  
 Joana Neca, 13  
 João A. Saraiva, 81  
 João Benedito, 289  
 João C. Silva, 81  
 João Faria, 245, 249  
 João M. Brisson Lopes, 39  
 João M. Fernandes, 121  
 João Morais, 293  
 João Oliveira, 291  
 João P. Ferreira, 141  
 João Pereira, 29, 129  
 João Teixeira, 155  
 Joaquim Jorge, 21, 129, 221, 237, 241, 289, 301  
 Jorge A. Santos, 141  
 Jorge Ferraz de Abreu, 283, 295  
 José Borges, 305  
 José Creissac Campos, 81, 121  
 José Luís Silva, 121  
 Júlia Teles, 113
- Leonor Teixeira, 73  
 Luís A. R. Neng, 271  
 Luís Alexandre, 205  
 Luís Bruno, 129  
 Luís Carriço, 13, 55, 59, 245, 249, 299  
 Luís Certo, 305  
 Luís Duarte, 55  
 Luís F. A. Silva, 189  
 Luís Miguel S Ponciano, 287  
 Luís Teixeira, 113
- Manuel J. Fonseca, 163, 167  
 Manuel Próspero dos Santos, 173  
 Marco de Sá, 55, 245, 249,  
 Maria Beatriz Carmo, 159  
 Maria Carreira, 281  
 Mário Rui Gome,s 65  
 Michael D. Harrison, 121  
 Miguel Sales Dias, 97, 285, 287
- Nádia Fernandes, 299  
 Nelson Baloian, 245  
 Nídia Salomé Morais, 197

Óscar Mealha, 263  
Óscar R. Ribeiro, 121  
Osvaldo Rocha Pacheco, 295

Paula Alexandra Silva, 303  
Paulo Dias, 47, 137, 141,  
Paulo Gomes, 147  
Paulo Ricardo Duarte, 285  
Paulo Santos, 137  
Pedro A. Lopes, 29  
Pedro Beça, 309  
Pedro Bernardo, 249  
Pedro Branco, 275  
Pedro Campos, 213  
Pedro Feiteira, 297  
Pedro Rangel Henriques, 89  
Pedro Ribeiro, 105

Ricardo Ameixa, 293  
Ricardo Dias, 163, 167  
Ricardo Jota, 21  
Ricardo Pereira, 245  
Rita Oliveira, 283  
Rita Pereira, 241  
Rui A. Costa, 47  
Rui Borges Lopes, 307

Rui José, 105  
Rui Lopes, 299  
Rui Nuno de Almeida, 97  
Rui Palha, 47  
Rui Raposo, 197, 309

Salvador Abreu, 205  
Samuel Almeida, 263  
Sandra Gama, 147, 163  
Shahram Izadi, 3  
Sofia Pessanha, 213

Tarquínio Mota, 33  
Telmo Silva, 295  
Teresa Chambel, 271  
Teresa G. Dias, 305  
Teresa Gama, 237  
Teresa Heitor, 281  
Teresa Romão, 193  
Tiago Gomes, 13  
Tiago Guerreiro, 221, 237, 241, 289, 291, 301  
Tiago Reis, 13, 59  
Tiago Silva, 105

Vasco Saavedra, 73  
Vitor Lago Silva, 275